Quantum Chemical Modeling of Binuclear Zinc Enzymes

Shilu Chen

Department of Theoretical Chemistry
School of Biotechnology
Royal Institute of Technology
Stockholm, Sweden, 2008
Abstract

In the present thesis, the reaction mechanisms of several di-zinc hydrolases have been explored using quantum chemical modeling of the enzyme active sites. The studied enzymes are phosphotriesterase (PTE), aminopeptidase from *Aeromonas proteolytica* (AAP), glyoxalase II (GlxII), and alkaline phosphatase (AP). All of them contain a binuclear divalent zinc core in the active site. The density functional theory (DFT) method B3LYP functional was employed in the investigations. The potential energy surfaces (PESs) for various reaction pathways have been mapped and the involved transition states and intermediates have been characterized. The hydrolyses of different types of substrates were examined, including phosphate esters (PTE and AP) and the substrates containing carbonyl group (AAP and GlxII). The roles of zinc ions and individual active-site residues were analyzed and general features of di-zinc enzymes have been characterized.

The bridging hydroxide stabilized by two zinc ions has been confirmed to be capable of the nucleophile in the hydrolysis reactions. PTE, AAP, and GlxII all employ the bridging hydroxide as the direct nucleophile. Furthermore, it is shown that either one of or both zinc ions provide the main catalytic power by stabilizing the negative charge developing during the reaction and thereby lowering the barriers. In the cases of GlxII and AP, one of zinc ions also contributes to the catalysis by stabilizing the leaving group. These features perfectly satisfy the two requisites for the hydrolysis, i.e. sufficient nucleophilicity and stabilization of charge. A competing mechanism, in which the bridging hydroxide acts as a base, was shown to have significantly higher barrier in the case of PTE.

For phosphate hydrolysis reactions, it is important to characterize the nature of the transition states involved in the reactions. Associative mechanisms were observed for both PTE and AP. The former uses a step-wise associative pathway via a penta-coordinated intermediate, while the latter proceeds through a concerted associative path via penta-coordinated transition states.

Finally, with PTE as a test case, systematic evaluation of the computational performance of the quantum chemical modeling approach has been performed. This assessment, coupled with other results of this thesis, provide an effective demonstration of the usefulness and powerfulness of quantum chemical active-site modeling in the exploration of enzyme reaction mechanisms and in the characterization of the transition states involved.
Acknowledgements

I have had a really good time in the past three years in which I have been working at Department of Theoretical Chemistry, Royal Institute of Technology, Stockholm. I would like to express my most intense gratitude to my supervisor, Dr. Fahmi Himo! Greatly appreciate everything he taught me and enjoy every fun we had together. This must be the most valuable and colorful three years in my life.

I also owe my sincere thanks to my supervisor in Beijing Normal University, Prof. Wei-Hai Fang. Thank him for his guidance and support during the last several years. The ten years in Beijing Normal University will be in my mind for ever.

Great thanks to Profs. Hans Ågren and Yi Luo, and thanks to everybody at Department of Theoretical Chemistry. I will forever miss everyone I met here and everything we enjoyed together.

A special thank is offered to our cooperator, Prof. Frank M. Rauschel, in Texas A&M University. I thank him for his valuable opinions and for the collaboration in the investigation of Phosphotriesterase.

I also greatly thank Profs. Nino Russo and Tiziana Marino, in the University of Calabria, for the collaboration on the study of the AAP mechanism.

Finally, endless thanks to my family who are always, always by my side!
List of papers included in this thesis

(I) **Theoretical Study of the Phosphotriesterase Reaction Mechanism**
Shi-Lu Chen, Wei-Hai Fang, Fahmi Himo

(II) **Technical Aspects of Quantum Chemical Modeling of Enzymatic Reactions: the Case of Phosphotriesterase**
Shi-Lu Chen, Wei-Hai Fang, Fahmi Himo

(III) **Structure of Diethyl Phosphate Bound to the Binuclear Metal Center of Phosphotriesterase**
Jungwook Kim, Ping-Chuan Tsai, Shi-Lu Chen, Fahmi Himo, Steven C. Almo, Frank M. Raushel
*Biochemistry, 2008, 47, 9497-9504.*

(IV) **Peptide Hydrolysis by the Binuclear Zinc Enzyme Aminopeptidase from *Aeromonas proteolytica*: A Density Functional Theory Study**
Shi-Lu Chen, Tiziana Marino, Wei-Hai Fang, Nino Russo, Fahmi Himo

(V) **Reaction Mechanism of the Binuclear Zinc Enzyme Glyoxalase II – A Theoretical Study**
Shi-Lu Chen, Wei-Hai Fang, Fahmi Himo

(VI) **Insights into the Transition States and Mechanism of Alkaline Phosphatase Reaction from the DFT Calculations**
Shi-Lu Chen, Wei-Hai Fang, Fahmi Himo
*Manuscript.*
List of papers not included in this thesis

(VII) Insights into Mechanistic Photodissociation of Acetyl Chloride by ab Initio Calculations and Molecular Dynamics Simulations
Shi-Lu Chen, Wei-Hai Fang

(VIII) Insights into Photodissociation Dynamics of Propionyl Chloride from ab Initio Calculations and Molecular Dynamics Simulations
Shi-Lv Chen, Wei-Hai Fang

(IX) Electrochemiluminescence of Terbium (III)-Two Fluoroquinolones-Sodium Sulfite System in Aqueous Solution
Shi-Lv Chen, Fen Ding, Yu Liu, Hui-Chun Zhao
*Spectrochimica Acta Part A*, 2006, 64, 130.

(X) Determination of Norfloxacin Using a Terbium-Sensitized Electrogeneated Chemiluminescence Method
Shi-Lv Chen, Yu Liu, Hui-Chun Zhao, Lin-Pei Jin, Zhong-Lun Zhang, Yan-Zhen Zheng
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>Aminopeptidase from <em>Aeromonas proteolytica</em></td>
</tr>
<tr>
<td>AP</td>
<td>Alkaline phosphatase</td>
</tr>
<tr>
<td>CPCM</td>
<td>Conductor-like Polarizable Continuum Model</td>
</tr>
<tr>
<td>DFT</td>
<td>Density functional theory</td>
</tr>
<tr>
<td>DHO</td>
<td>Dihydroorotase</td>
</tr>
<tr>
<td>GlxII</td>
<td>Glyoxalase II</td>
</tr>
<tr>
<td>HF</td>
<td>Hartree-Fock</td>
</tr>
<tr>
<td>LFER</td>
<td>Linear free energy relationship</td>
</tr>
<tr>
<td>MFJ plot</td>
<td>More O’Ferral Jencks plot</td>
</tr>
<tr>
<td>MO</td>
<td>Molecular orbital</td>
</tr>
<tr>
<td>PCM</td>
<td>Polarizable Continuum Model</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PES</td>
<td>Potential energy surface</td>
</tr>
<tr>
<td>PSI</td>
<td>Phosphoserine intermediate</td>
</tr>
<tr>
<td>PTE</td>
<td>Phosphotriesterase</td>
</tr>
<tr>
<td>TS</td>
<td>Transition state</td>
</tr>
<tr>
<td>TST</td>
<td>Transition state theory</td>
</tr>
<tr>
<td>ZPE</td>
<td>Zero-point energy</td>
</tr>
</tbody>
</table>
### Amino acids abbreviations

<table>
<thead>
<tr>
<th>1-Letter symbol</th>
<th>3-Letter symbol</th>
<th>Full name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ala</td>
<td>Alanine</td>
</tr>
<tr>
<td>C</td>
<td>Cys</td>
<td>Cysteine</td>
</tr>
<tr>
<td>D</td>
<td>Asp</td>
<td>Aspartate</td>
</tr>
<tr>
<td>E</td>
<td>Glu</td>
<td>Glutamate</td>
</tr>
<tr>
<td>F</td>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>G</td>
<td>Gly</td>
<td>Glycine</td>
</tr>
<tr>
<td>H</td>
<td>His</td>
<td>Histidine</td>
</tr>
<tr>
<td>I</td>
<td>Ile</td>
<td>Isoleucine</td>
</tr>
<tr>
<td>K</td>
<td>Lys</td>
<td>Lysine</td>
</tr>
<tr>
<td>L</td>
<td>Leu</td>
<td>Leucine</td>
</tr>
<tr>
<td>M</td>
<td>Met</td>
<td>Methionine</td>
</tr>
<tr>
<td>N</td>
<td>Asn</td>
<td>Asparagine</td>
</tr>
<tr>
<td>P</td>
<td>Pro</td>
<td>Proline</td>
</tr>
<tr>
<td>Q</td>
<td>Gln</td>
<td>Glutamine</td>
</tr>
<tr>
<td>R</td>
<td>Arg</td>
<td>Arginine</td>
</tr>
<tr>
<td>S</td>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>T</td>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>V</td>
<td>Val</td>
<td>Valine</td>
</tr>
<tr>
<td>W</td>
<td>Trp</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>X</td>
<td>Xaa</td>
<td>Any residue</td>
</tr>
<tr>
<td>Y</td>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
</tbody>
</table>
Contents

Chapter 1 Introduction .................................................................................................................. 1

Chapter 2 Theoretical background ............................................................................................ 3
  2.1 Wave function methods ....................................................................................................... 3
  2.2 Density functional theory ................................................................................................. 5
  2.3 Performance of B3LYP ..................................................................................................... 7
    2.3.1 B3LYP accuracy on geometry ................................................................................. 8
    2.3.2 B3LYP accuracy on energy ..................................................................................... 8
  2.4 Deficiencies of DFT ......................................................................................................... 10

Chapter 3 Enzyme catalysis and its modeling ............................................................................. 11
  3.1 Enzyme catalysis .............................................................................................................. 11
  3.2 Transition state theory ..................................................................................................... 14
  3.3 Modeling of enzymatic reactions ..................................................................................... 14
    3.3.1 Construction of active site model ........................................................................... 15
    3.3.2 Computational methods .......................................................................................... 16
  3.4 Evaluation of modeling approach with PTE as a test case ................................................ 17
    3.4.1 Effect of basis set .................................................................................................... 18
    3.4.2 Choice of dielectric constant .................................................................................. 20
    3.4.3 Effect of locking atoms ........................................................................................... 21
    3.4.4 Summary ................................................................................................................ 22

Chapter 4 General features of di-zinc enzymes ........................................................................ 23
  4.1 Nature of bridging species ............................................................................................... 23
  4.2 Nucleophile vs. base mechanisms .................................................................................... 25
  4.3 Associative, dissociative, and $S_N2$ mechanisms ............................................................. 26
  4.4 Suggested roles of Zn ion ................................................................................................. 28
    4.4.1 Binding and orienting the substrate ....................................................................... 28
    4.4.2 Lower $pK_a$ of the hydrolytic water molecule ....................................................... 28
    4.4.3 Stabilizing charge developing during reaction ....................................................... 28
    4.4.4 Other functions of binuclear metal center ............................................................... 29
  4.5 Roles of other amino acids in the active site ..................................................................... 30

Chapter 5 Applications .............................................................................................................. 31
  5.1 Phosphotriesterase (PTE) (Papers I and III) ..................................................................... 31
    5.1.1 Structure and mechanism ....................................................................................... 31
    5.1.2 Nucleophile mechanism for dimethyl 4-nitrophenyl phosphate substrate .......... 32
    5.1.3 Base mechanism for dimethyl 4-nitrophenyl phosphate substrate ....................... 34
    5.1.4 Trimethyl phosphate substrate ............................................................................... 35
    5.1.5 Conclusions ............................................................................................................ 36
  5.2 Aminopeptidase from Aeromonas proteolytica (AAP) (Paper IV) ................................. 36
    5.2.1 Structure and mechanism ....................................................................................... 37
    5.2.2 Active site model .................................................................................................... 37
5.2.3 Nucleophilic attack .................................................................................................................. 38
5.2.4 Proton transfer ........................................................................................................................... 39
5.2.5 C-N bond cleavage ....................................................................................................................... 40
5.2.5 Conclusions ................................................................................................................................. 41
5.3 Glyoxalase II (GlxII) (Paper V) .................................................................................................... 43
  5.3.1 Structure and mechanism ........................................................................................................ 43
  5.3.2 Active site model ........................................................................................................................ 44
  5.3.3 Nucleophilic attack ..................................................................................................................... 45
  5.3.4 C-S Bond cleavage ..................................................................................................................... 47
  5.3.5 Product release and active site regeneration ............................................................................ 47
  5.3.6 Conclusions ............................................................................................................................... 50
5.4 Alkaline Phosphatase (AP) (Paper VI) ....................................................................................... 51
  5.4.1 Structure and mechanism ..................................................................................................... 51
  5.4.2 Active site model ...................................................................................................................... 53
  5.4.3 Methyl phosphate substrate .................................................................................................... 55
  5.4.4 p-Nitrophenyl phosphate substrate ........................................................................................ 57
  5.4.5 Hydrolysis of phosphoseryl intermediate ............................................................................... 58
  5.4.6 Conclusions ............................................................................................................................. 60
Chapter 6 Conclusions .................................................................................................................... 62
References ............................................................................................................................................ 63
Chapter 1
Introduction

Enzymes are proteins that work as the indispensable catalysts for the various reactions in the organisms. In human body, for example, a large number of known or unknown enzymes catalyze the innumerable reactions in order to sustain diverse critical biological functions, such as metabolism, perception, motion, cell regulation, and so on. The investigation of enzymatic reactions is thus a crucial key to discover the numerous mysteries of life, like infection, immunity, drug resistance, and consciousness. The objects of this thesis are di-zinc enzymes, which contain two divalent zinc ions in their active sites. All known di-zinc enzymes are hydrolases and function in different biochemical events, such as protein maturation and degradation, tissue repair, physiological detoxification, and cell-cycle control.

Zinc is silent for most of spectroscopic techniques, a fact that often hampers experimental mechanistic investigations of zinc enzymes. Although the zinc ion can often be replaced by other metal ions (especially Co$^{2+}$ and Cu$^{2+}$) without the loss of the catalytic activity, particular care obviously needs to be paid. One must therefore rely on the help of modern quantum chemical techniques to explore the reaction mechanisms of zinc enzymes. In recent years, advancements in density functional theory (DFT) methods, in particular the development of the hybrid B3LYP functional, coupled with the constant growth of computer power, have made it possible to treat ever larger systems at a reasonable level of accuracy. Using these methods, one can today routinely handle systems containing more than 100 atoms, a development that has made it possible to study enzymatic reactions. One very fruitful approach has been to cut out a relatively small model around the active site and evaluate it by high-level quantum mechanics, while the effects of the missing protein surrounding are estimated by crude approximations, such as locking atoms and using polarizable continuum model method. Researchers have used this approach to successfully investigate mechanistic aspects of a wide range of different enzymes.

In the present thesis, we used this approach to investigate the reaction mechanisms of several di-zinc hydrolases. The studied enzymes are phosphotriesterase (PTE), aminopeptidase from Aeromonas proteolytica (AAP), glyoxalase II (GlxII), and alkaline phosphatase (AP). The hydrolyses of different types of substrates were examined, including phosphate esters (PTE and AP) and the substrates containing carbonyl group (AAP and GlxII). The roles of zinc ions and individual active-site residues were analyzed and some general features of di-zinc enzymes have been characterized. In particular, some important but disputed issues were scrutinized, like for instance the nucleophilicity of bridging hydroxide, nucleophile vs. base mechanisms, and associative vs. dissociative mechanisms.

In this thesis, a brief outline of theoretical background (Chapter 2) and
computational approach (Chapter 3) will be presented before the results of theoretical studies of enzyme reactions. An evaluation of the computational approach with PTE as a test case will also be discussed in Chapter 3. Some general features of di-zinc enzymes will be discussed in Chapter 4, followed by the studies of individual enzymes in Chapter 5. Finally, some conclusions will be made in Chapter 6.
Chapter 2
Theoretical background

Theoretical chemistry is the field where mathematical methods are combined with fundamental laws of physics to study chemical problems. Quantum mechanics has been the most powerful theory used to describe the microscopic chemistry, in particular the motion of electrons in a molecule. The general equation in quantum mechanics is the time-dependent Schrödinger equation, by solving which we can in principle obtain all properties of a system, especially the energy. For the ground-state chemistry, like the enzymatic reactions studied in the present thesis, it is sufficient to use the time-independent form of the Schrödinger equation \((\hat{H}\Psi = E\Psi)\) as the base. The exact solution of the Schrödinger equation encounters the great difficulty. The only systems that can be solved exactly are those composed of only one or two particles. To calculate the many-body systems, a number of approximations are necessary.

2.1 Wave function methods

All properties of a system are in principle derived from the motion of electrons and nuclei in the molecules, which can be generally described by the Schrödinger equation. The time-independent form of Schrödinger equation \((\hat{H}\Psi = E\Psi)\) is usually sufficient to be utilized for the ground-state chemistry\(^{(1,2)}\). For a general N-particle system, the Hamilton operator \((\hat{H})\) contains kinetic \((\hat{T})\) and potential energy \((\hat{V})\) for all nuclei (denoted by “n”) and electrons (e):

\[
\hat{H} = \hat{T} + \hat{V} = \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ne} + \hat{V}_{ee}
\]

\[
= -\sum_k \frac{1}{2M_k} \nabla_k^2 - \sum_i \frac{1}{2} \nabla_i^2 + \sum_{k=1}^N \frac{Z_kZ_L}{|\mathbf{R}_k - \mathbf{R}_j|} - \sum_{i,k} \frac{Z_k}{|\mathbf{r}_i - \mathbf{R}_k|} + \sum_{i>j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}
\]

Where \(\nabla_i^2 = \left( \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right)\)

For the many-particle systems, i.e. beyond \(H_2^+\), this equation is too complicated
to be solved exactly. Therefore, some simplification have to be introduced. The most fundamental approximation, known as Born-Oppenheimer approximation, is based on the fact that nuclei move much slower than the electrons, and their motions can be separated. The kinetic energy of nuclei can thus be ignored and the nucleus-nucleus repulsion is constant for certain geometry. The Schrödinger equation can thus be separated into two parts which describe the nuclear and electronic wave functions, respectively. The solution to the electronic Schrödinger equation, i.e. calculating the electronic energies for different nuclear arrangements, leads to a potential energy surface (PES), the minima of which determine the equilibrium geometries of a molecule.

For a many-body system, the electronic Schrödinger equation is still too complicated to be solved, so additional approximations are needed. One of the most important approximations is the Hartree-Fock (HF) method, which is based on the independent-particle model, where each electron is considered to move in the mean field of all other particles. Each electron is thus related with a one-electron wave function (called molecular orbital, MO), which is the combination of a spatial function that depends on the coordinate of the electron, and a spin function that depends on its spin. The wave function has to satisfy the antisymmetry principle, and must change sign if the coordinates of two electrons are interchanged (Pauli principle). A basic way to build the wave function is by using a Slater determinant of the N one-electron orbitals (N is the number of electrons). With this, the N-particle problem is transformed to a set of one-particle problems:

$$\hat{f}_i \chi_i = \epsilon_i \chi_i$$  \hspace{1cm} (2.2)

where $\hat{f}_i$ is an effective one-electron operator, in which the electron-electron repulsion is treated in an average way. $\chi_i$ is the corresponding eigenfunction (i.e. MO) and the electron in the MO has the orbital energy $\epsilon_i$. HF equation is non-linear and has to be solved in iterative approach, the procedure of which is called the self-consistent field (SCF) method. The orbitals in a single Slater determinant can then be optimized to minimize the energy. In practice, the molecular orbitals in a molecule are usually constructed as a linear combination of the atomic orbitals (LCAO) of the corresponding atoms.

The independent-particle model results in an inherent limitation of the HF method, since the motion of all electrons are correlated in a real system. Neglecting correlation energy\footnote{The difference between the HF energy and the exact non-relativistic ground-state energy within the Born-Oppenheimer approximation.} leads to large deviations from experimental results, which makes the HF method very poor for exploring chemical reactions. A number of approaches to correct this weakness, collectively called post-Hartree-Fock methods\cite{3}, have been...
developed to include electron correlation. For example, Møller-Plesset perturbation theory treats correlation as a perturbation of the Fock operator. Multi-configurational self-consistent field (MCSCF), configuration interaction (CI), coupled cluster (CC) and complete active space SCF (CASSCF) expand the true multi-electron wave function in terms of a linear combination of Slater determinants. These approaches improve the level of accuracy but become computationally much more demanding, and thus are only suitable for relatively small systems at the moment. To investigate large systems, as the enzymatic reactions in this thesis, a “cheaper” alternative method is needed.

2.2 Density functional theory

A very fruitful alternative to the wave function methods is the density functional theory (DFT) approach\textsuperscript{[2,4,5]}. The basis of DFT is the Hohenberg-Kohn theorem\textsuperscript{[6]}, which shows that the total energy of a non-degenerate ground state is a unique functional of the electron density of the system, namely, \( E = E[\rho(r)] \). This implies that all properties of a system can be deduced from the ground-state density and the determination of the complicated many-electron wave function can thus be avoided. However, a fundamental difficulty emerges here, that is, the exact functional, i.e. the dependency of the energy on the given electron density, is not known. Various approximations and attempts have been made.

The energy functional can be expressed as follows\textsuperscript{[2,4,5]}:

\[
E[\rho] = T[\rho] + E_{ee}[\rho] + E_{ne}[\rho]
\] (2.3)

where \( T \) is the kinetic energy, \( E_{ee} \) the electron-electron repulsion, and \( E_{ne} \) the nuclei-electron attraction. In 1965, Kohn and Sham\textsuperscript{[7]} contributed a significant development, i.e. the orbital-based scheme, in which independent particles move in an effective potential (the non-interacting one electron orbitals are called Kohn-Sham orbitals, \( \phi_i \)). The real system of interacting electrons can thus be described through a system of non-interacting particles by expressing the electron density as the sum of the squared orbitals. Therefore, the total kinetic energy (\( T \)) is divided into two parts, the kinetic energy of an N electrons non-interacting system (\( T_s \)) and a missing fraction (\( T_c \)) relative to the real interacting system:

\[
T[\rho] = T_s[\rho] + T_c[\rho]
\] (2.4)

The functional of electron-electron repulsion (\( E_{ee}[\rho] \)) can be divided into the
classical Coulomb interaction \((J)\) and a non-classical part containing correlation and exchange \((E_{ncl}[\rho])\):

\[
E_{xc}[\rho] = J[\rho] + E_{ncl}[\rho]
\]  

(2.5)

The total energy can now be written as:

\[
E[\rho] = T_s[\rho] + T_c[\rho] + J[\rho] + E_{ncl}[\rho] + E_{nc}[\rho]
\]  

(2.6)

Then, a definition is done by combining the missing part of the kinetic energy \((T_c[\rho])\) and the correlation and exchange part \((E_{ncl}[\rho])\) to form an exchange-correlation functional \((E_{xc}[\rho])\). The total energy can finally be presented as:

\[
E[\rho] = T_s[\rho] + J[\rho] + E_{nc}[\rho] + E_{xc}[\rho]
\]  

(2.6)

The first three terms can be calculated explicitly. All problems have now been centralized in how to accurately describe the exchange-correlation term, \(E_{xc}[\rho]\), which incorporates all unknown contributions to the total energy.

Using the variational principle together with the normalization constraints, minimizing the total energy of a determinant constructed by Kohn-Sham orbitals results in the Kohn-Sham equations (similar to the Hartree-Fock equation):

\[
\hat{h}_{ks} \phi_i(r) = \varepsilon_i \phi_i(r)
\]  

(2.7)

where \(\hat{h}_{ks}\) is the one-electron operator and depends on the electron density. If the exact form of the \(E_{xc}[\rho]\) functional is known, the exact total energy of the many-electron system can be obtained by iteratively solving the equation (2.7). The accuracy of a DFT method lies on how accurate THE form of \(E_{xc}[\rho]\) is.

Many exchange and correlation functionals have been or are being developed. A significant improvement to the accuracy of DFT came from the introduction of the gradient of the electron density in the functional, i.e. \(E_{xc}[\rho, \nabla \rho]\). Later, Becke’s
2. Theoretical background

introduction of the exact Hartree-Fock exchange as a part of $E_{xc}^{\rho, \nabla \rho}$\textsuperscript{[8,9]} successfully leads to the popularity of DFT. The DFT methods including HF exchange are referred to as hybrid methods. The predominant hybrid functional used by chemists is the B3LYP functional, which has been employed in the present thesis. The B3LYP functional is written as a linear combination of HF exchange and local- and gradient-corrected exchange and correlation:

$$E_{xc}^{\text{B3LYP}} = aE_{x}^{HF} + (1-a)E_{x}^{\text{Slater}} + bE_{x}^{\text{B88}} + cE_{c}^{\text{LYP}} + (1-c)E_{c}^{\text{VWN}}$$ \hspace{1cm} (2.8)

The weighting parameter $a$ determines the extent of replacement of the Slater local exchange ($E_{x}^{\text{Slater}}$) by the exact HF exchange ($E_{x}^{HF}$); $b$ controls the addition of Becke’s gradient-correction to the exchange functional ($E_{x}^{\text{B88}}$)\textsuperscript{[10]}; $c$ defines the inclusion weight of the LYP correlation ($E_{c}^{\text{LYP}}$)\textsuperscript{[11]} and the VWN correlation ($E_{c}^{\text{VWN}}$)\textsuperscript{[12]} functionals\textsuperscript{1}. The three coefficients were optimized by minimizing the average absolute deviation of theory from experiment for 116 atomic and molecular properties (56 atomization energies, 42 ionization potentials, 8 proton affinities, and 10 first-row total atomic energies)\textsuperscript{[9]}.

2.3 Performance of B3LYP

An important advantage of the DFT methods relative to the wave-function-based methods is the lower scaling. For the DFT methods, the dependency of computational time ($t$) on the number of basis functions ($N$) is $t \sim N^\alpha$ ($\alpha \approx 2$-3)\textsuperscript{3}, while $\alpha$ is larger in the wave-function methods (for HF, $\alpha = 4$). The relatively low computational cost makes the DFT methods possible to be applied to the large systems. However, the employment of the DFT methods is definitely affected by its accuracy, in particular the accuracy on geometry and energy. To test the computational performance of the DFT methods, especially the B3LYP functional, various benchmark tests have been performed. Here the B3LYP accuracy on geometries and energies will be discussed briefly. Within this discussion, the requirement of the basis set size for the geometry optimization and energy evaluation will also be mentioned.

\begin{itemize}
  \item[1] LYP is a gradient-corrected correlation functional, while VWN is an electron-gas correlation functional.
  \item[2] It can also be approximatively considered as the number of atoms
  \item[3] For so-called linear scaling methods, $\alpha$ can even be 1.
\end{itemize}
2.3.1 B3LYP accuracy on geometry

The accuracy of various DFT methods with respect to geometry and energy has been tested on the standard G2 benchmark test, composed of 55 small first- and second-row molecules\(^{[13]}\). In this test, the computational performance on structural parameters, involving 71 bond lengths, 26 bond angles, and 2 dihedral angels, were evaluated (summarized in Table 2.1). The most important conclusion from Table 2.1 is that all DFT methods give quite accurate geometries, already with a medium-sized basis set. For example, the average unsigned error at the B3LYP/6-31G(d) level are 0.013 Å for bond lengths, 0.62 Å for angles, and 0.35 degree for dihedral angels. For the B3LYP functional, the application of a larger basis set, 6-311+G(3df,2p), just slightly reduces the errors in bond lengths and angels (see Table 2.1). The error in dihedral angles increases (3.66°) when the larger basis set was used. This is, however, not a reliable result since only two dihedral angles were tested. Furthermore, if the geometries obtained using the larger basis set were employed for the atomization energy evaluation, instead of the 6-31G(d) geometries, the average deviation only varies from 2.22 to 2.20 kcal/mol\(^{[13]}\). These results indicate that B3LYP has good accuracy on geometries and that a medium-sized basis set is generally sufficient for the geometry optimization, a finding that is also applicable to the modeling of enzyme reactions in this thesis. A special evaluation concerning the effect of basis set on the geometries in the reaction modeling of the di-zinc-containing enzymes (phosphotriesterase as the example), will be presented in Section 3.4 and Paper II.

<table>
<thead>
<tr>
<th>Methods(^a)</th>
<th>Bond lengths (Å)</th>
<th>Angles (deg)</th>
<th>Dihedral angles (deg)</th>
<th>Energies (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>0.020</td>
<td>1.16</td>
<td>1.92</td>
<td>74.50</td>
</tr>
<tr>
<td>MP2</td>
<td>0.015</td>
<td>0.67</td>
<td>1.24</td>
<td>7.43</td>
</tr>
<tr>
<td>BP</td>
<td>0.026</td>
<td>1.03</td>
<td>0.89</td>
<td>11.81</td>
</tr>
<tr>
<td>BLYP</td>
<td>0.020</td>
<td>0.91</td>
<td>0.27</td>
<td>4.95</td>
</tr>
<tr>
<td>BP86</td>
<td>0.022</td>
<td>0.96</td>
<td>0.24</td>
<td>10.32</td>
</tr>
<tr>
<td>B3P86</td>
<td>0.010</td>
<td>0.62</td>
<td>0.86</td>
<td>7.82</td>
</tr>
<tr>
<td>B3LYP</td>
<td>0.013</td>
<td>0.62</td>
<td>0.35</td>
<td>2.22</td>
</tr>
<tr>
<td>B3LYP(big)(^b)</td>
<td>0.008</td>
<td>0.61</td>
<td>3.66</td>
<td>2.20</td>
</tr>
</tbody>
</table>

\(^a\) The 6-31G(d) was used for geometry optimization and the 6-311+G(3df,2p) basis set for the energies.
\(^b\) The 6-311+G(3df,2p) basis set was used for geometry optimization and energies.

2.3.2 B3LYP accuracy on energy

Many tests have been made to examine the B3LYP accuracy with respect to various energies, including atomization energy, ionization potential, electron affinity, proton affinity, barrier height, and so on.

Bauschlicher’s test for the G2 set (55 molecules) (Table 2.1) shows that the
B3LYP method is clearly superior to the other methods with respect to atomization energy, with a mean error of 2.20 kcal/mol at the B3LYP/6-311+G(3df,2p) level\(^\text{[13]}\). Quite recently, a more extensive evaluation was carried out by Curtiss et al.\(^\text{[14]}\) to test a number of density functional methods\(^\text{1}\) on the G3/05 test set. This set includes 454 energies\(^\text{2}\), all of which have experimental uncertainties less than ±1 kcal/mol. The assessment shows that B3LYP performs well with a mean unsigned deviation of 4.14 kcal/mol. It has high-accuracy performance for some energies such as proton affinities and hydrogen-bond strengths (the mean unsigned deviations are 1.39 and 1.19 kcal/mol, respectively), whereas it works less well for enthalpies of formation (4.63 kcal/mol).

Since the investigation in the present thesis focuses on the reaction mechanisms of enzymes, how the B3LYP method performs with respect to barriers is especially important. In recent years, a number of examinations have been carried out in this aspect. The ability of 15 density functionals to calculate barrier heights was evaluated by Zhao and Truhlar\(^\text{[15]}\) for a dataset of 42 reactions, comprising mainly open-shell hydrogen-transfer processes. B3LYP gave a mean unsigned error of 4.31 kcal/mol. Zhao and Truhlar\(^\text{[16]}\) have also analyzed the behavior of 25 density functionals in the calculation of 38 barrier heights for non-hydrogen-transfer reactions, including 12 heavy-atom-transfer reactions, 16 nucleophilic substitution reactions, and 10 non-nucleophilic unimolecular association reactions. B3LYP performed quite well, with an average unsigned deviation of 3.08 kcal/mol for the total 76 barrier heights\(^3\). However, it was shown to systematically underestimate the barriers, particularly in heavy-atom-transfer reactions (the mean signed error is -8.49 kcal/mol). A recent study by Riley et al.\(^\text{[17]}\) has compared 37 density functionals in the calculation of barrier heights for 23 reactions of small systems with radical transition states. In this investigation, B3LYP ran with a mean unsigned error of 4.30 kcal/mol. Riley’s study also examined the performance in the barrier determination for six reactions of larger systems with singlet transition states, resulting in a mean unsigned error of 3.10 kcal/mol for B3LYP. The overall tests above indicate that the average error of B3LYP on reaction barriers for various reactions appear to be a few kilocalories per mole. This level of accuracy makes B3LYP quite reliable to be utilized in the exploration of reaction mechanisms. Unfortunately, for enzymatic reactions, no extensive benchmarks have been done. Siegbahn concludes that B3LYP in general gives the error of around 3 kcal/mol in relative energies of enzymatic reactions for molecules containing first- and second-row atoms\(^\text{[18]}\). For systems involving transition metals, the error appears to be somehow larger, but rarely more than 5 kcal/mol\(^\text{[18]}\).

As the theme of this thesis is related to the enzymes containing two zinc ions, it is of interest to evaluate the performance of B3LYP for Zn coordination chemistry. Quite recently, Amin and Truhlar\(^\text{[19]}\) tested the predictions of 39 density functionals in 10 Zn-ligand bond distances, 8 dipole moments, and 12 bond dissociation energies of

---

\(^1\) Totally 23 methods including a number of new functionals, such as X3LYP, O3LYP, TPSS, and so on.

\(^2\) In particular some energies of hydrogen-bonded complexes and molecules containing third-row elements were included.

\(^3\) The mean unsigned errors are 8.49 kcal/mol for heavy-atom-transfer reactions, 3.25 kcal/mol for nucleophilic substitution, and 2.02 kcal/mol for non-nucleophilic unimolecular association.
Zn coordination compounds with O, S, NH₃, H₂O, OH, SCH₃, and H ligands. B3LYP performed well in the aspects of Zn-ligand bond distance and dipole moment (the mean unsigned errors are 0.0080 Å and 0.44 D, respectively), but less well in bond dissociation energy (5.41 kcal/mol).

It is worth being stressed here that the basis-set dependence of B3LYP for energies is significant. In most of benchmarks tests, increasing the basis set would result in significant improvement in the energy accuracy when using B3LYP. Combined with the basis-set dependence of B3LYP for geometries, a common procedure employed in this thesis emerges, i.e. optimizing the geometry using a smaller basis set and then evaluating the energy using a quite large basis set. Usually, the 6-311+G(2d,2p) basis set is required for the final energy evaluation.

2.4 Deficiencies of DFT

The recent development of DFT has made it an efficient and popular tool in computational chemistry. However, it is important to remember that DFT is not exact or perfect yet. Several deficiencies exist in DFT indeed, mainly involving self-interaction errors, near-degeneracy errors, and the lack of description of Van der Waals interactions.

In wave-function methods (like Hartree-Fock), the artificial repulsion between an electron and itself is exactly canceled by an exchange term. In DFT, Coulombic terms are described exactly, but the exchange is described by an approximate functional. These terms do not exactly cancel in DFT, leading to so-called self-interaction error[5]. This error artificially stabilizes delocalized transition states and tends to decrease the barrier heights.

The second error, called near-degeneracy error, is due to the inherent description of the wave function as a single determinant (the non-dynamical correlation is lacking)[18]. In contrast to the effect of the self-interaction error, this error trends to increase barrier heights. Therefore, there is a substantial cancellation effect between self-interaction and near-degeneracy errors. In practice, the B3LYP functional appears to be built to balance these errors as well as possible. For the high-electron-delocalization reactions, such as hydrogen atom, or proton, transfer reaction, the barriers are usually underestimated since the self-interaction error prevails in these cases. While the barriers of most other kinds of reactions often trends to be overestimated as the near-degeneracy error predominates.

The third deficiency of DFT is the lack of a description of Van der Waals interactions. This deficiency often leads to exaggerated repulsion when atoms are forced close to each other, which usually happens in the systems with several large substituents or ligands.
Chapter 3
Enzyme catalysis and its modeling

Enzymes are bio-macromolecules that serve as the catalysts to accelerate chemical reactions in the organisms. Almost all enzymes are proteins and range in size probably from just 62 amino acid residues for the monomer of 4-oxalocrotonate tautomerase\(^{[20]}\) to over 2500 residues for the animal fatty acid synthase\(^{[21]}\). Most of enzymes are highly efficient and often increase the reaction rate thousands of times relative to the corresponding reaction in the aqueous solution. Some enzymes are very selective, not only in choice of substrates with different conformations but also in the stereochemistry of the reactions. These features make enzymes indispensable for life. Thus, it is conceivably very important to understand, in detail, how the enzymes work. In this situation, the modeling of enzymatic reactions has become a very valuable tool in this pursuit.

3.1 Enzyme catalysis

The reason why enzymes can speed up reactions is that they can reduce the activation barriers of the reactions, which determine the reaction rates\(^{[22,23]}\). An illustration of this is presented in Figure 3.1. For the uncatalyzed reaction (A), the reaction barrier, \(\Delta G^*\), is the relative free energy of the transition state (TS) to the reactant (substrate, S). In the corresponding enzymatic reaction (B), the substrate first binds to the enzyme to form an enzyme-substrate complex (ES), and then react via one (or several) transition state (TS\(_{cat}\)) resulting in an enzyme-product complex (EP), from which the product (P) is finally released and the enzyme is set free again. In this case the reaction barrier (\(\Delta G_{cat}^*\)) is the relative energy between the ES complex and the transition state\(^1\). The enzyme can provide the catalytic power to make \(\Delta G_{cat}^* < \Delta G^*\), and consequently accelerates the rate of the overall reaction. The catalytic power can be thought to originate mainly from the stabilization of the transition state, and in some cases from straining the shape of the substrate into its transition state form, i.e. ground state destabilization\(^2\). Hence, a crucial mission in the investigation of enzymatic reaction is to find out how the enzyme can stabilize its transition state more than the transition state of the uncatalyzed reaction. The possibly

---

\(^1\) For the multiple-step enzymatic reaction, the overall reaction barrier is the largest energy difference between an intermediate (the ES complex should, of course, be taken into account) and a transition state.

\(^2\) Both of them can reduce the amount of energy for the transition.
most effective way for achieving large stabilization is the use of electrostatic effects[24]. An example of this is the use of a cation cofactor to stabilize the developing negative charge of the transition state. Another way to achieve stabilization is to hold a relatively fixed polar environment that is orientated toward the charge distribution of the transition state. The former is a general strategy employed by the di-zinc enzymes studied in this thesis. In addition, many other proposals have been put forward to explain the great catalytic activity of enzymes, such as (i) desolvation hypothesis (the nonpolar environment inside the enzyme destabilize the highly charged reacting state), (ii) entropic guidance (enzymatic reaction takes place with the ES complex as the starting point and hardly costs the translational and rotational entropy; the binding energy pays the cost of the entropy), (iii) high effective concentration (a high effective concentration of the participating groups can be achieved by the binding of the substrate in the active site), (iv) orbital steering (the orbitals of the reacting atoms or molecules are aligned in the suited positions), (v) dynamic effects (the motions of enzyme may lead to the lowering of activation barrier and the fluctuation of the transmission coefficient; the former result is expected to have greater contribution to the catalytic activity, since the barrier is in the exponential of the rate expression while the transmission coefficient a prefactor), (vi) tunneling (usually occurs in the light-particle transfer reactions, in particular electron and proton transfers), and so on. Many of these ideas are highly controversial. More detailed discussions about the different proposals can be found in the references[24-28].

![Figure 3.1 Schematic free energy profile for an uncatalyzed (A) and an enzymatic (B) reaction.](image)

For a multiple-step enzymatic reaction with several intermediates of various stability, the largest energy difference between an intermediate and a transition state in the forward direction corresponds to the rate-limiting step. The overall rate of an enzyme-catalyzed reaction can be described by a rate constant $k_{cat}$. The $k_{cat}$ is then equivalent to the rate constant of the rate-limiting step during the reaction. Thus, one can calculate the overall rate constant $k_{cat}$ using classical transition state theory (see
3. Enzyme catalysis and its modeling

Section 3.2) if the rate-limiting barrier has been obtained by computational approach. This is of significance since this theoretical rate constant of an enzyme is related to the experimental maximum rate (called Turnover Rate), which can often be experimentally determined at saturating substrate concentration.

The free energy of the overall reaction \( \Delta G^0 \) is the energy difference between the reactant \((S)\) and product \((P)\). It is independent of the enzyme (see Figure 3.1) because, in the enzyme-catalyzed reaction, the same product is formed and the enzyme is retained unchanged after the reaction. This energy \( \Delta G^0 \) is also called the Driving Force of the reaction. One can easily obtain the driving force of an enzymatic reaction by experimentally determining or theoretically calculating the energy of the corresponding uncatalyzed reaction. The energy difference between \(E+S\) and the \(ES\) complex shows the binding energy (see \(a\) in Figure 3.1), while the difference between the \(EP\) complex and \(E+P\) is the release energy (\(c\) in Figure 3.1). In Figure 3.1, both these energies are indicated to be negative. This is, however, not necessarily always the case. In principle, the enzyme should not bind the substrate too tightly, because the formation of a deep energy minimum of the \(ES\) complex would decrease the efficiency of the enzyme. The binding is therefore achieved basically by intermolecular forces, such as salt bridge, hydrogen bond, Van der Waals force, and hydrophobic interaction. The binding energy gives rise to the specificity of an enzyme by energetically distinguishing various substrates. Different enzymes have different specificity: some have the very high specificity as they only catalyze one substrate, while others exhibit catalytic promiscuity in that they can act on a relatively broad range of substrates. However, it is rather difficult to accurately determine the binding energy using the quantum chemical methods employed in the present thesis. To obtain accurate binding energy requires sufficiently large model of the active site and also very accurate bulk solvent representation. If the energy of reaction pathway (see \(b\) in Figure 3.1) has already been obtained, one can roughly estimate the sum of binding energy and release energy by calculating the difference between overall reaction energy and reaction energy, i.e. \(a + c = \Delta G^0 - b\). This concept has been applied in the investigation of the active-site regeneration of Glyoxalase II (see Section 5.3 and Paper V).

In an enzyme-catalyzed reaction, there could be several different reaction pathways leading to the same product. Quantum chemical methods can be very useful in discriminating between the different mechanisms, by means of a potential energy surface (PES) and by identifying intermediates (minima) and transition states. The calculated energies are many times enough to support or discard a suggested mechanism. Furthermore, an advantage of quantum chemical approach is the ability to optimize and characterized short-lived species that cannot be detected by most experimental techniques. These capabilities make quantum chemical methods a powerful tool in studying the enzymatic reaction mechanisms.
3.2 Transition state theory

Classical transition state theory (TST) provides a very simple but powerful way to connect the calculated energies to the measured rate constants. The rate constant, according to TST, can be determined using\cite{2}

\[ k = \frac{k_B T}{h} e^{-\frac{\Delta G^*}{RT}} \]  

(3.1)

where \( k_B \) is Boltzmann’s constant and \( h \) is Planck’s constant. At room temperature (298.15 K), a rate of 1 s\(^{-1}\), i.e. one unit of reaction per second, can thus be calculated to correspond to a barrier of 17.4 kcal/mol. Also, a change of 1.4 kcal/mol in barrier means an approximate raise or fall in reaction rate by one order of magnitude. These are very useful relationships to remember when assessing the feasibilities of some calculated reaction mechanisms.

3.3 Modeling of enzymatic reactions

In recent years, quantum chemical methods have had very positive impacts on the study of enzymatic reaction mechanisms. One very fruitful approach to explore the reaction mechanisms has been to cut out a relatively small model of enzyme, where the reaction takes place, and optimize it at a quite accurate level of theory (B3LYP is used in the present thesis). The part of the enzyme that is not included in the quantum model can be modeled by means of two main approximations. The ignored enzyme surrounding has steric effect on the active site, forcing various groups to stay in certain positions or preventing them from making free and unorderly movements. Thus, the first approximation is to fix certain atoms in the model (typically where the truncations are made) to their X-ray positions. In addition, to account for the polarization effects, the enzyme surrounding is approximately considered as a homogeneous polarizable medium which can be modeled using the dielectric cavity techniques. This approach has previously been successfully used to investigate the reaction mechanisms of various enzymes\cite{29-33}. In this section, this methodology of enzyme modeling will be discussed briefly.
Figure 3.2 Construction of an active site model of an enzyme, the case of aminopeptidase from Aeromonas proteolytica (AAP). (Upper) Overall X-ray crystal structure of AAP and close-up view of the active site. Coordinates from PDB entry 1RTQ were used to generate the figures. (Lower) the chemical model of AAP active site. Stars indicate the atoms that are fixed to their X-ray positions. The solvation effect is simulated using the polarizable continuum methods.

3.3.1 Construction of active site model

In the modeling of enzyme reaction, the prerequisite structural information is directly taken from the available X-ray crystal structure. The residues and cofactors that are proposed to be important for a suggested reaction mechanism are extracted from the PDB file. For example, in the building of the active site model of aminopeptidase from Aeromonas proteolytica (see Figure 3.2), the first-shell residues, the zinc ion cofactors, a bridging hydroxide, and a second-shell glutamic acid (Glu151,
proposed to act as a proton shuttle) are taken from the crystal structure. To reduce the size of the model, the residues are truncated such that in principle only the functional groups of the amino acids are kept in the model. However, to gain more flexibility for the active site, one more carbon of the side chain is usually kept in the model. In the example of Figure 3.2, it is better to represent the histidines by methyl-imidazoles, rather than imidazoles. Likewise, acetate is a better model for the aspartate or glutamate than formate. In some cases, even more carbons must be retained in the model, especially for residues that make big movements during the reaction. Then, the hydrogen atoms are added manually according to the chemical knowledge and the expected protonation states of the residues. The substrate of interest is finally introduced into the model. The binding position and orientation of the substrate is determined on the basis of the crystal structure of enzyme in complex with substrate, substrate analogue, or/and inhibitor. In some instances, different substrate orientations have to be tested to evaluate whether the orientation is critical for the mechanism.

3.3.2 Computational methods

The DFT functional B3LYP is employed in all calculations presented in this thesis. As pointed out above, using a medium-sized basis set can obtain sufficient accuracy on geometry in the B3LYP calculation. Therefore, all geometry optimizations (including minima and transition states) were carried out with the 6-31G(d,p) basis set for all elements except the transition metal (Zn), for which the effective core potential LANL2DZ basis set was used. To account for the steric effect of the missing surrounding, some atoms where truncations have been done are fixed in the optimizations to preserve their spatial arrangements. Frequency calculations were performed at the same theory level as the optimizations to obtain zero-point energies (ZPE) and to confirm the nature of the stationary points. This implies no negative eigenvalues for minima and only one negative eigenvalue for transition states. The procedure of locking atoms can give rise to a few imaginary frequencies, typically on the order of $20i$ cm$^{-1}$. These frequencies do not contribute significantly to the ZPE.

The energies obtained from the small-basis set optimizations are not accurate enough for the evaluation of mechanism, so some corrections should be added. On the basis of the optimized structures, more accurate energies can be achieved by performing single-point calculations with the larger basis set 6-311+G(2d,2p). Then, the zero-point energies from frequency calculations should be included. Since computing frequencies is very time-consuming, for very large quantum model, the ZPE effects from a smaller model describing the same reaction are sometimes used.

The solvation effect provided by the protein surrounding should also be added. In our cases, the surrounding is simply considered as a homogeneous polarizable medium, which can be simulated using polarizable continuum model (PCM) methods with a certain dielectric constant ($\varepsilon$). Therefore, the solvation effects can be easily estimated at the same theory level as the optimizations by performing the single-point
calculations on the optimized structure using some PCM methods. In practice, the CPCM\cite{34} and IEF-PCM\cite{35} methods were frequently used. The dielectric constant of the surrounding medium was typically chosen to be 4, which corresponds roughly to a mixture of a proteinaceous medium ($\varepsilon = 2-3$) and water ($\varepsilon = 80$)\cite{36}. Several tests have been carried out to assess the influence of different dielectric constants on the energetics\cite{37,38}. In this thesis, an evaluation has also been done on the choice of dielectric constant in the modeling of the di-zinc enzyme reactions (see next section and Paper II).

Finally, all calculations in this thesis were performed using the Gaussian03 program package\cite{39}.

### 3.4 Evaluation of modeling approach with PTE as a test case

The modeling approach presented above has previously been successfully applied to the studies of the reaction mechanisms of a number of enzymes\cite{29-33}. However, it is still interesting to know how well the approach works with respect to the di-zinc enzymes. To achieve this, we performed a series of calculations to evaluate the adequacy of the technical approximations used in the modeling of the di-zinc enzyme reactions. As a test case, we chose phosphotriesterase (PTE), a di-zinc enzyme that catalyzes the hydrolysis of organophosphate triesters. The content of this section is a summary of Paper II.

PTE is a homodimer with each subunit containing a di-zinc active site\cite{40,41}. The two zinc ions are connected to the protein matrix via the side chains of His55, His57, His201, His230, and Asp301. In addition, a hydroxide and a carboxylated lysine residue (Lys169) bridge the two zinc ions. A reaction mechanism for PTE has been proposed as follows\cite{42,43}: the bridging hydroxide (O$_\mu$H) performs a nucleophilic attack on the substrate from its bridging position to form a penta-coordinated intermediate, followed by the leaving group departing without activation (protonation). We have used the quantum chemical approach outlined above to study the reaction of PTE with a model consisting of 82 atoms (shown in Figure 3.3). In this study, we did not fix the atoms to their crystallographic positions where the truncation was made, since the zinc ions could anchor the various parts of the active site, and also because no second-shell residues were included in the model. The calculations give strong support to this mechanism and suggest that the second step is the rate-limiting step (see Section 5.1 and Paper I). In this section, we will not discuss the mechanism, but rather focus on the technical issues involved in the calculation of the potential energy profile for the enzyme. These include the choice of basis set for geometry optimizations and energy evaluation, the choice of dielectric constant, and the effects of locking the centers of truncation.
3. Enzyme catalysis and its modeling

Figure 3.3 Model of the PTE active site used in the investigation. Coordinates were taken from the crystal structure (PDB code 1HZY). Key bond distances are defined.

3.4.1 Effect of basis set

We have first optimized the geometries using the LANL2DZ basis set on the zinc ions and the 6-31G(d,p) on all other atoms. This will be called Basis Set 1 (BS1). The most important geometric parameters of the stationary points along the reaction pathway using BS1 are listed in Table 3.1 (see the detailed structures in Paper I). It can be first found that the overall geometric parameters of the reactant (React) agree very well with the X-ray structure. For example, the distances of the bridging oxygen to the two zinc ions are calculated to be 2.00 and 1.99 Å, to be compared to the crystallographic distances of 2.0 Å.

Table 3.1 Important geometric parameters for the various stationary points using different models.

<table>
<thead>
<tr>
<th></th>
<th>React</th>
<th>TS1</th>
<th>Inter</th>
<th>TS2</th>
<th>Prod</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BS1</td>
<td>BS2</td>
<td>BS1</td>
<td>BS2</td>
<td>BS1</td>
</tr>
<tr>
<td>r1</td>
<td>3.54</td>
<td>3.53</td>
<td>3.41</td>
<td>3.69</td>
<td>3.71</td>
</tr>
<tr>
<td>r2</td>
<td>2.00</td>
<td>2.00</td>
<td>1.97</td>
<td>2.15</td>
<td>2.14</td>
</tr>
<tr>
<td>r3</td>
<td>1.99</td>
<td>1.98</td>
<td>2.01</td>
<td>2.22</td>
<td>2.22</td>
</tr>
<tr>
<td>r4</td>
<td>2.24</td>
<td>2.28</td>
<td>2.20</td>
<td>2.04</td>
<td>2.02</td>
</tr>
<tr>
<td>r5</td>
<td>3.42</td>
<td>3.45</td>
<td>3.00</td>
<td>2.24</td>
<td>2.12</td>
</tr>
<tr>
<td>r6</td>
<td>1.63</td>
<td>1.62</td>
<td>1.64</td>
<td>1.75</td>
<td>1.74</td>
</tr>
<tr>
<td>r7</td>
<td>1.50</td>
<td>1.48</td>
<td>1.50</td>
<td>1.55</td>
<td>1.53</td>
</tr>
<tr>
<td>r8</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>1.03</td>
<td>1.02</td>
</tr>
<tr>
<td>r9</td>
<td>1.79</td>
<td>1.81</td>
<td>1.86</td>
<td>1.51</td>
<td>1.56</td>
</tr>
<tr>
<td>r10</td>
<td>2.14</td>
<td>2.14</td>
<td>2.16</td>
<td>2.11</td>
<td>2.11</td>
</tr>
</tbody>
</table>

The parameters are defined in Figure 3.3. BS1: LANL2DZ on Zn and 6-31G(d,p) on all other atoms. BS2: Same as BS1 but using 6-311+G(2d) on the phosphorous center and the five oxygens around it. F: The model in which the truncation atoms are fixed.
The resulting potential energy profile using BS1 is shown in Figure 3.4. The nucleophilic attack by the O$_{μ}$H$^{-}$ on the substrate is calculated to have a barrier of 8.2 kcal/mol (TS1) and the resulting penta-coordinated intermediate (Inter) lies at +5.2 kcal/mol relative to the React. From Inter, the barrier for the departure of the leaving group (TS2) is 5.4 kcal/mol (+10.6 kcal/mol relative to React) and the resulting product (Prod) is 0.8 kcal/mol lower than React. When applying a larger basis set (LBS: 6-311+G(2d,2p) on all atoms) as a single-point energy calculation on the optimized structures, some significant changes can be observed on the PES (see Figure 3.4). For example, the energies of TS1, Inter, and TS2 are considerably raised relative to React, while the energy of Prod is lowered. This indicates that the small basis set (BS1) is not entirely adequate for a proper description of the changes during the reaction, and thus a larger basis set is needed to obtain reliable energetics.

This raises the question whether a larger basis set is also required to obtain reliable geometries. To investigate this, all the geometries have been re-optimized using the same basis set as before (BS1), but with the important difference that for the phosphorus center and all five oxygens around it the 6-311+G(2d) was used. This combination is called Basis Set 2 (BS2). The new PES using this basis set is displayed in Figure 3.4, and the geometric parameters are listed in Table 3.1. We can find some small changes in the geometries around the phosphor center. The largest changes are seen for the forming O$_{μ}$-P bond ($r_5$ in Figure 3.3) in TS1, and the breaking P-O$_{L}$ bond ($r_6$) in TS2, both of which are shortened by ca 0.1 Å.

The better description of the phosphorus center and its surrounding atoms leads to a considerable difference in the energetics compared to BS1. The energies are now much closer to the ones obtained by LBS//BS1 than those obtained with BS1 (Figure
However, a very important result here is that when applying the LBS on the BS2-optimized geometries, the PES is very similar to the one obtained by applying the LBS on the BS1-optimized geometries. The differences are less than 1 kcal/mol. There is thus no advantage in using a larger basis set for the geometry optimizations.

In summary, the results here demonstrate that geometries can be optimized using a medium-sized basis set, while the final energies should definitely be evaluated using a much larger basis set.

### 3.4.2 Choice of dielectric constant

The choice of the dielectric constant for the surrounding medium is somewhat arbitrary, but the value of \( \varepsilon = 4 \) is often used. To investigate what effect the choice of the dielectric constant has on the energetics, we have used two different values, \( \varepsilon = 4 \) and \( \varepsilon = 80 \). The solvation effects are calculated as single-point corrections at the same level of theory as the geometry optimizations. We have done this using both the BS1- and BS2-optimized geometries and added the effect to the energies from the large basis set calculations. The results are shown in **Figure 3.5**.

![Figure 3.5](image)

**Figure 3.5** Calculated potential energy profile for the PTE reaction using different dielectric constants. Left: Geometries and solvation effects calculated with BS1. Right: Geometries and solvation effects calculated with BS2.

As seen from **Figure 3.5**, solvent effects do not change the energetics of the first step significantly. In the second step, however, the effects are more pronounced. It is easy to rationalize these results considering that the P-O bond cleavage of the second step results in an anionic phenolate product that dissociates from the active site. Upon addition of solvation, the energy of this species should be lowered considerably. Using BS1- and BS2-geometries gave the similar results.

Another finding is that the solvation effects saturate very quickly as a function of the dielectric constant. In going from the cluster model (\( \varepsilon = 1 \)) to \( \varepsilon = 4 \), most of solvation has been accounted for. The difference between \( \varepsilon = 4 \) and \( \varepsilon = 80 \) is not very large.
3. Enzyme catalysis and its modeling

Most importantly, the choice of the dielectric constant does not change any conclusion about the mechanism. These results indicate the exact choice of the dielectric constant is less critical in the quantum chemical modeling of enzyme reactions. Furthermore, a general conclusion from the calculations is that unless charges are created or quenched close to the edge of the quantum model, the relative solvent effects are usually quite small. This is because the most of the polarization has already been included into the cluster model.

3.4.3 Effect of locking atoms

As pointed out above, a common procedure in the modeling of enzymatic reactions is to fix some atoms during the optimization typically where the truncation is made. To examine how much this kind of procedure affects the relative energies of the reaction, we have re-done all the calculations using the same model of Figure 3.3, but with the truncation centers fixed to their X-ray positions. Six atoms were kept fixed, as indicated by the stars in Figure 3.6. The geometrical parameters are summarized in Table 3.1 and the resulting PES is shown in Figure 3.7.

Some significant differences can be observed between the constrained and unconstrained geometries. For example, the Zn-Zn distance ($r_1$ in Figure 3.3 and Table 3.1) for the React, Inter, and Prod structures is 3.54, 3.69, and 4.34 Å, respectively, for the unconstrained model, to be compared to 3.41, 3.46, and 3.59 Å, respectively, for the constrained model. This indicates that the strain is increasing from React to Prod. This strain is indeed reflected in the computed energies: Prod is now 0.4 kcal/mol higher than React, compared to -4.6 kcal/mol in the unconstrained model. In addition, the distance of the bridging oxygen to the phosphor ($r_5$) is shortened to 3.00 Å from 3.42 Å in the unconstrained model. The differences in the other stationary points are smaller, which suggests that, in the constrained model, React is more strained and destabilized relative to the other stationary points. As a result, the energies of TS1, Inter, and TS2 are all shifted down compared to React.

Figure 3.6 Optimized structure of the reactant complex with constraints. Stars indicate the atoms that are fixed to their X-ray positions. All distances are in angstrom (Å).
Figure 3.7 Potential energy curves for the PTE reaction with and without constraints on the truncation atoms. Both of curves were calculated at the 6-311+G(2d,2p)//BS1 level.

Although the procedure of locking the truncation atoms to their X-ray positions seems to result in significant strain, as revealed in both the geometries and the potential energy curves, these changes are not such of a magnitude that they will alter any conclusions about the mechanism of PTE. For example, the second step can be stated to be the rate-limiting step no matter which model were used.

It should be added here that usually one more carbons are kept in the amino acid side chains, which can lead to less pronounced effects of the locking scheme.

3.4.4 Summary

Based on the calculations above, it can be concluded that geometries can be calculated using a medium-sized basis set without a significant loss of accuracy. However, the final energy should be evaluated using a quite large basis set. Moreover, the choice of the dielectric constant used to model the protein surrounding is less critical. The solvation effects are usually quite significant for steps that create or quench charge at the edge of the quantum model, for example, like the case of the departure of the charged leaving group. Although the way of locking the atoms where the truncation is made can lead to a significant amount of strain if the model is not large enough, this approximation was not found to change the conclusions drawn about the reaction mechanism. In summary, the approach outlined above (Section 3.3) is sufficiently reliable to be utilized in studying the reaction mechanisms of di-zinc enzymes.
Chapter 4

General features of di-zinc enzymes

Zinc is an extremely important metal biologically as it is present in many proteins, including enzymes and DNA binding proteins. Although zinc’s abundance inside human is relatively low, its concentration in liver, voluntary muscle, and bone reaches 60-180 μg/g\[44\]. Zinc usually exists in the oxidation state of +2, i.e. in the form of Zn\(^{2+}\) ions. Some properties of zinc ion, such as flexibility in coordination geometry, fast ligand exchange, Lewis acidity, intermediate polarization and/or stabilization, are important and useful in its function in enzymes. Zn\(^{2+}\) can interact with a wide range of biological ligands, such as His, Asp, Glu, Cys, amino group of peptide, and so on, and typically has a tetrahedral or trigonal bipyramidal arrangement of its ligands when bound to enzymes. In a few cases, Zn\(^{2+}\) can be hexa-coordinated (for example, see Glyoxalase II in Section 5.3 and Paper V).

In living organisms, innumerable hydrolysis reactions have to be performed efficiently, in order to sustain a large number of critical biological functions. In some cases nature has taken care of this elegantly by placing two (or even three) divalent metal ions (Zn\(^{2+}\), Mg\(^{2+}\), Ni\(^{2+}\), Mn\(^{2+}\), Co\(^{2+}\), Cu\(^{2+}\), Fe\(^{2+}\)) close to each other (M-M distance = 3-5 Å)\[45,46\] in the active site of an enzyme. In particular, di-zinc core is frequently employed in the active site. The two ions work together to bind and activate the substrate and a water molecule, which then undergo hydrolysis under biological conditions (room temperature and neutral pH, in which these reactions normally cannot occur without the enzymes). Removal of one or both metal ions from the active site usually causes a significant decrease in activity or total inactivation, indicating that both ions work synergistically as a single unit. Thus, a series of interesting issues emerge for binuclear metallo-hydrolases, e.g. what are the reaction mechanisms; what about the roles of metal ions; is there any common principle for the function of this type of enzymes. In this chapter some of the general features of binuclear zinc enzymes will be discussed.

4.1 Nature of bridging species

The two metal ions in most of di-zinc enzymes are found to be bridged by a molecule from solvent, which possibly is water (H\(_2\)O), hydroxide (OH\(^{-}\)), or even oxide (O\(^{2-}\)) (see Scheme 4.1)\[47\]. In most of cases, this bridging species is supposed to act as the key nucleophile in the hydrolysis. Therefore, its identification is absolutely necessary.
Coordination of water molecule to two zinc ions will significantly lower its pK$_a$ value, which makes the bridging species quite possible to be hydroxide. To further identify the form of the bridging species, a number of experimental and theoretical studies have been performed. Using PTE as the example, a kinetic experiment reported that the pK$_a$ of Zn/Zn-PTE was 5.8$^{[62]}$, while molecular simulation on the thermodynamics of protonation of the bridging hydroxide gave a value of 5.1$^{[48]}$. These pK$_a$ values point out that hydroxide is the bridging species in Zn/Zn-PTE. The active site of Mn/Mn-PTE has also been investigated by utilizing X-band EPR spectroscopy$^{[49]}$. The Mn/Mn-PTE binuclear center exhibits an exchange coupling constant$^1$ of $J = -2.7 \pm 0.2$ cm$^{-1}$, indicating that a hydroxide bridges the two metal ions. Moreover, most of other di-zinc enzymes containing the bridging species are proposed to use the hydroxide as the bridge$^{[47,50]}$. All di-zinc enzymes studied in this thesis, except for AP, also employ the hydroxide as the bridge.

In most of cases, the bridging hydroxide (O$_\mu$H$^{-}$) stabilized by two metal ions is suggested to function as the crucial nucleophile in the hydrolysis. However, it has been questioned that the nucleophilicity of bridging hydroxide is not sufficient to be utilized in the hydrolysis reaction$^{[51]}$. Some studies have suggested that the hydroxide has to transform to the terminal from the bridging state before it can perform the nucleophilic attack$^{[52,53]}$. In this thesis we have tested the nucleophilicity of the bridging hydroxide for a number of di-zinc enzymes, including PTE, AAP, and GlxII. As we shall see below, our calculations strongly demonstrate that the nucleophilicity of the bridging hydroxide is sufficient to be utilized in the hydrolysis reaction (see Sections 5.1, 5.2, and 5.3). This is consistent with the fact that protonation of the bridging hydroxide results in the loss of catalytic activity of PTE$^{[49]}$. The low catalytic efficiency of the synthesized binuclear zinc complexes in the biomimetic modeling$^{[51]}$ probably attributes to the poor stabilization of the developing negative charge and improper orientation of the substrate.

It is widely accepted that the main reason why most of hydrolysis reactions are very slow in solution is that the solvent water molecule is a bad nucleophile. The di-zinc enzymes, however, have solved this problem very elegantly. Once a water molecule has been bound by binuclear zinc center of an enzyme, it can be immediately converted into a hydroxide, a much better nucleophile. The resulting hydroxide will subsequently be located at a perfect position for the step of nucleophilic attack. This partially explains why the hydrolysis reactions catalyzed by

1 Magnetic coupling of metal centers is often facilitated by their bridging ligands through a super-exchange mechanism. The sign and magnitude of the exchange coupling constant ($J$), are influenced by the metal ions, bridging ligands, terminal ligands, metal-bridge-metal angle, metal-metal distance, and metal-bridge distance.
the di-zinc enzymes are much faster than the corresponding reactions in solution.

4.2 Nucleophile vs. base mechanisms

In most of cases, the bridging hydroxide is suggested to directly function as the nucleophile in the hydrolysis, which is referred to as “nucleophile mechanism” (see Scheme 4.2). Quite recently, however, Ollis and co-workers have proposed a “base mechanism” for the PTE reaction[54]. The OµH⁻ acts not as a nucleophile but as a base to abstract a proton from a water molecule that is loosely coordinated to the more-buried α-metal ion (see Scheme 4.2). Apparently, we can get two significant questions, i.e. (i) which mechanism is preferable for PTE reaction, and (ii) whether base mechanism is available for other di-zinc enzymes. Actually, this kind of base mechanism has also been advised for the reactions of alkylsulfatase from *Pseudomonas aeruginosa*[55] and purple acid phosphatase[56] on the basis of the crystal structures.

To identify the different mechanisms for PTE, Raushel et al. have recently solved two structures of PTE complexed with diethyl phosphate (DEP, the hydrolysis product) and cacodylate (a product analogue), respectively (see Paper III). In the structure of the PTE-DEP complex, the DEP product is found symmetrically bridging the two divalent cations and no bridging hydroxide is detected. A similar structure is observed in the PTE-cacodylate complex. These results strongly support the nucleophile mechanism that the bridging hydroxide directly performs the nucleophilic attack and finally ends up at the phosphate product.

![Nucleophile mechanism](image1)

![Base mechanism](image2)

*Scheme 4.2 The nucleophile and base mechanisms in PTE reaction.*

To further examine the feasibility of the base mechanism proposed by Ollis et al., we have carried out the DFT calculations using a model of the Zn/Zn-PTE active site. Our calculations show that a hexa-coordinated α-Zn is quite unlikely and the base
mechanism is energetically less favorable than the nucleophile mechanism (see Section 5.1 and Paper III). The nucleophile mechanism thus turns out to be preferable for the PTE reaction.

4.3 Associative, dissociative, and $S_N2$ mechanisms

Phosphate ester hydrolysis is a critically important reaction in biological systems, involved in signal and energy transduction, biosynthesis, protein synthesis, and replication of the genetic material\[^{57}\]. In this thesis, we also study examples of the phosphate ester hydrolysis, namely the reactions catalyzed by two di-zinc enzymes, phosphotriesterase (PTE, Paper I, II, and III) and alkaline phosphatase (AP, Paper VI). They catalyze the hydrolysis of phosphotriesters and phosphate monoesters, respectively. It is here necessary to generally introduce the reaction mechanisms of phosphate ester hydrolysis. As a model, we take the phosphate monoester hydrolysis with a hydroxide as the nucleophile. A free energy surface for the hydrolysis of phosphate monoester can be represented as a More O’Ferral Jencks (MFJ) plot\[^{58}\] (see Figure 4.1).

![Figure 4.1 More O’Ferral Jencks (MFJ) diagram for the phosphate monoester hydrolysis. (O$_{\text{nuc}}$) The nucleophile oxygen. (O$_{\text{lg}}$) The leaving-group oxygen.](image)

---

- 26 -
Attack of hydroxide on the phosphorus center will generally proceed through one of three different pathways, i.e. associative, dissociative, and SN2 pathways. In an associative SN1-type mechanism, the nucleophile attacks prior to leaving group departure and the reaction proceeds via a penta-coordinated intermediate (the species in the top left corner of Figure 4.1). In contrast, in a dissociative mechanism, the leaving group departure precedes the nucleophilic attack and between them a three-coordinated metaphosphate intermediate (the bottom right state) is formed. Both associative and dissociative pathways are thus step-wise. There is, however, a third possibility, a concerted SN2-type mechanism, in which bond formation and bond cleavage occur in a single reaction step (see the diagonal pathways in Figure 4.1). This reaction passes via a single transition state without an intermediate. A concerted transition state may be either synchronous (with roughly equal amounts of bond formation and bond cleavage, see TS_{syn} in Figure 4.1) or asynchronous. The latter can further be either associative (with a larger degree of bond formation to the nucleophile than breaking of the bond to the leaving group, TS_{ass}) or dissociative (with more bond dissociation than bond formation, TS_{diss}).

In view of the complexity of phosphate ester hydrolysis, it is difficult, but very interesting, to study the reaction mechanisms of enzymatic hydrolysis, in particular to characterize the transition states. A well-known experimental method for probing reaction mechanisms is the use of linear free energy relationships (LRERs)\[59\]. This approach involves studying the hydrolysis of a series of homologous compounds with various pKa values. The logarithm of the obtained rate constant is correlated to the pKa of the leaving group such that:

\[
\ln k_{cat} = \beta pK_a + C
\]

(4.1)

The Bronsted coefficient, \(\beta\), determines the leaving group dependence of the reaction series. A reaction proceeding through a dissociative mechanism would be expected to have a large negative value of \(\beta\), while the associative mechanism would yield a smaller \(\beta\) since the reaction is expected to be less dependent on the nature of the leaving group. This approach is effective only when the chemical step is rate-limiting and the substituent-specific effect is minor. It has a number of shortcomings: (i) LFERs cannot discriminate between concerted and step-wise mechanisms; (ii) LFERs are not based on direct molecular information. It has been demonstrated that different mechanistic possibilities can lead to similar LFERs\[60\]. For example, the associative and dissociative pathways may yield the same LFER. This indicates that the experimental LFERs do not have unique mechanistic interpretations\[60\].

In this situation, the exact location of the transition states along reaction pathway using quantum chemical methods (the approach has been discussed in Chapter 3) can be a powerful method to explore the mechanisms of enzymatic reactions. Using quantum chemical modeling methods, we have investigated the reactions of PTE (Section 5.1 and Paper I) and AP (Section 5.4 and Paper VI) in this thesis. The associative pathway has been detected for both PTE and AP reactions.
4.4 Suggested roles of Zn ion

For the investigation of the di-zinc enzymes, it is crucial to find what the roles of zinc ions are during the catalytic reaction. This is important for the understanding of reaction mechanism and enzyme evolution. Based on previous experimental studies and our calculations, the main roles of zinc ions in the di-zinc enzymes can be summarized as below.

4.4.1 Binding and orienting the substrate

As a divalent cation, zinc ion can play a role in binding and orientating the substrate to be in an excellent position for the reaction, in particular for the nucleophilic attack. Let us take AAP as the example: an optimized structure of the AAP active site with the model substrate bound has been shown in Figure 3.2. It can be concluded, from our analysis (see details in Section 5.2.2), that one role of the Zn2 center is to bind the substrate, assisting in orienting the peptide bond toward the nucleophile (the bridging hydroxide). This is consistent with the fact that the AAP enzyme exhibits 80% of its activity upon removal of only one metal ion, while the complete removal of the two ions inactivates it entirely.

4.4.2 Lowering the pKa of the hydrolytic water molecule

As discussed in Section 4.1, coordination of water molecule to metal ion can lower its pK_a value. The coordination to one metal ion can lower its pK_a from 15.7 in bulk solvent to as low as 9.0^{[61]}. If water is coordinated to two metal ions, its pK_a can be further decreased to ~ 6.0^{[62]}. The certain amount of lowering depends on the metal ions, the oxidation state of ions, other ligands, and the local dielectric. The pK_a lowering of water makes the hydroxide to be the most possible bridge. All di-zinc enzymes studied in this thesis, except for AP, are suggested to use the hydroxide as the bridge.

4.4.3 Stabilizing charge developing during reaction

The nucleophilic attack on the carbonyl carbon (C_C) or phosphorus atom (P) will develop negative charge at the carbonyl oxygen (O_C) or phosphoryl oxygen (O_P). For the reaction to take place, this developing charge should be stabilized. In all cases we studied, the zinc ions, which coordinate the O_C or O_P, were found to play this decisive role. This conclusion can be made by following the distance changes of C_C-O_C and Zn-O_C (or P-O_P and Zn-OP) in the nucleophilic attack step. Table 4.1 has
listed the distances of $C_C-O_C$ and $Zn-O_C$ (or $P-O_P$ and $Zn-O_P$) in the nucleophilic attack step for some di-zinc enzymes studied in this thesis.

From Table 4.1, we can find that the $C_C-O_C$ (or $P-O_P$) distances are elongated in going from the reactant to the intermediate via the transition state, indicating that the negative charge is developed at the $O_C$ (or $O_P$) atom. As a result, the distances of $Zn-O_C$ (or $Zn-O_P$) become shorter. This strongly demonstrates that the zinc ions play a role in stabilizing the developing charge. Consequently, the stabilization of the intermediate (or product) structures lowers the barriers of reactions. Without the stabilization of this kind, the catalysis would fail. Therefore, we can say that zinc ions provide the main catalytic power by stabilizing the charge developing during the reactions of di-zinc enzymes.

In some reactions, like GlxII (Section 5.3 and Paper V) and AP (Section 5.4 and Paper VI), one of zinc ion has another important role in stabilizing the negatively charged leaving group. This kind of stabilization lowers the barrier of the bond cleavage step and greatly contributes to the catalysis.

<table>
<thead>
<tr>
<th>Enzymes</th>
<th>$C_C-O_C$</th>
<th>$Zn-O_C$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reactant</td>
<td>TS</td>
</tr>
<tr>
<td>AAP</td>
<td>1.23</td>
<td>1.30</td>
</tr>
<tr>
<td>GlxII</td>
<td>1.22</td>
<td>1.27</td>
</tr>
<tr>
<td>PTE</td>
<td>1.50</td>
<td>1.52</td>
</tr>
</tbody>
</table>

*a Distance are given in angstrom. *b For PTE, the $P-O_P$ distances are given instead. *c For PTE, the $Zn-O_P$ distances are given instead.

4.4.4 Other functions of binuclear metal center

A very interesting issue is what the coupled functions of two metal ions are. Furthermore, what are the advantages of a binuclear metal center over a mononuclear center? A rough summary on the advantages is made here: (i) larger charge delocalization, (ii) smaller activation barriers due to delocalized charge, (iii) the ability to bind larger substrates, and (iv) easier ionization of water molecule. A detailed comparison between mononuclear and binuclear metal-binding sites can be found in a quite recent study, which combines a Protein Data Bank survey of binuclear sites with the DFT calculations[63]. However, it is not always true that the catalytic efficiency of the binuclear metal enzymes is greater than that of the mononuclear form. For example, for the binuclear zinc-β-lactamases from *Bacillus cereus* and *Bacteroides fragilis*, it has been shown that the catalytic activity is lower relative to the mononuclear form with some substrates[64]. Actually, the issues discussed in this section are still open questions.
4.5 Roles of other amino acids in the active site

In all cases containing the bridging hydroxide we studied, an amino acid with carboxylate group (aspartate or glutamate) is always found to make a hydrogen bond to the O$_\mu$H$^\cdot$. For example, in the crystal structure of PTE in complex with an inhibitor (diisopropyl methylphosphonate, DIMP), an aspartate (Asp301) is situated very close to the bridging oxygen (O$_\mu$) and suggested to make a hydrogen bond to the O$_\mu$ (see Figure 4.2). This residue is expected to play a role in orienting the hydroxide at the proper position to perform the nucleophilic attack. As seen from Figure 4.2, the lone pair points to the phosphorus center of the inhibitor by the help of the hydrogen bond. Mutation of the Asp301 to alanine or asparagine will lead to significant decrease in the catalytic activity of PTE$^{[42]}$. Similar experimental results can also be found in the cases of AAP$^{[65]}$ and GlxII$^{[66]}$.

Furthermore, this residue hydrogen bonding to the O$_\mu$H$^\cdot$ has another role in some cases, i.e. shuttling the proton of O$_\mu$H$^\cdot$. In the case of AAP, the second-shell Glu151 makes a hydrogen bond to the O$_\mu$H$^\cdot$ and is proposed to transport the proton to the nitrogen of the substrate during the reaction (see Section 5.2 and Paper IV). Without this function, the enzyme becomes inactive. In PTE reaction, the proton of the O$_\mu$H$^\cdot$ is not required for the activation of the leaving group. Instead, the Asp301 residue, coupled with the His254 and Asp233, can ferry away the proton from the active site$^{[42]}$.

In some cases, such as isoaspartyl dipeptidase and alkaline phosphatase, an amino acid (tyrosine and arginine are frequently used) can be observed to make the hydrogen bond to the carbonyl oxygen (or phosphoryl oxygen) of the substrate. This amino acid partially contributes to the catalysis by supplementarily stabilizing the developing negative charge at the O$_\nu$ (or O$_\gamma$) atom.

---

Figure 4.2 Relative orientation of the bridging hydroxide and the inhibitor DIMP (diisopropyl methylphosphonate) in the bacterial Zn/Zn-PTE (PDB Code: 1EZ2). The proposed orientation of the hydrogen and lone pair on the bridging hydroxide are shown for discussion purpose only.
Chapter 5
Applications

Using the quantum chemical approach outlined in Section 3.3, we have studied the reaction mechanisms of several di-zinc enzymes. These are phosphotriesterase (PTE), aminopeptidase from *Aeromonas proteolytica* (AAP), glyoxalase II (GlxII), and alkaline phosphatase (AP). The results will be summarized in this chapter.

5.1 Phosphotriesterase (PTE) (Papers I and III)

Phosphotriesterase (PTE) is a bacterial binuclear zinc enzyme that catalyzes the hydrolysis of a wide range of organophosphate triesters\(^\text{[67]}\), some of which are employed as agricultural insecticides and chemical warfare nerve agents. PTE is a very efficient enzyme, with a turnover number close to \(10^4 \text{ s}^{-1}\) for the best substrates, and a \(k_{\text{cat}}/K_{\text{m}}\) value of \(10^8 \text{ M}^{-1}\text{s}^{-1}\), which is close to the diffusion-controlled limit\(^\text{[68]}\).

5.1.1 Structure and mechanism

The X-ray crystal structure of PTE revealed that two zinc ions are bridged by a hydroxide from the solvent and a carboxylated lysine residue\(^\text{[69]}\). Four histidines, an aspartate, and a water molecule complete the first coordination sphere around the zinc ions. Three distinct binding pockets at the active site help orienting the substrate. It has been shown that the reaction proceeds with inversion of the stereochemical configuration of the phosphorus center and that the water oxygen is found in the phosphate product\(^\text{[70]}\).

Quite recently, however, Ollis and co-workers proposed a “base mechanism” for the PTE reaction (see Scheme 4.2)\(^\text{[54]}\), where the \(O_\mu H^-\) acts as a base to abstract a proton from a water molecule that is coordinated to the more-buried \(\alpha\)-metal ion. In this case, the \(\alpha\)-metal ion achieves the infrequent hexa-coordinated configuration, and
the water molecule serves as the nucleophile instead of the $\text{O}_\mu \text{H}^-$.

### 5.1.2 Nucleophile mechanism for dimethyl 4-nitrophenyl phosphate substrate

We used the high resolution crystal structure (PDB code 1HZY) to construct a model of the active site of PTE. This model is the same as the one used in Section 3.4 and has been shown in Figure 3.3. It consists of the two zinc ions and their first shell ligands, including the bridging hydroxide, the four histidines (His-55, His-57, His-201, His-230), Asp-301, and the carboxylated Lys-169. Hydrogen atoms were added manually and the ligands were truncated as shown in Figure 3.3. The total charge of the model is +1. The substrate used in the investigation is dimethyl 4-nitrophenyl phosphate, which was bound to the more solvent-exposed $\beta$-Zn site. The overall geometric parameters obtained from the geometry optimization of the present model of the PTE active site agree very well with the X-ray structure (see Paper I for the detailed structural parameters).

![Figure 5.1](image)

**Figure 5.1** Optimized structures of the transition state for the bridging hydroxide attack on the phosphorus (TS1), the resulting penta-coordinated intermediate (Inter), the following transition state for the P-O$_\lambda$ bond cleavage (TS2), and the resulting enzyme-product complex (Prod).

The transition state for the nucleophilic attack by the bridging hydroxide on the substrate (TS1, Figure 5.1) was optimized and the calculated energetic barrier for this step is the very feasible 10.8 and 10.5 kcal/mol, for $\epsilon = 4$ and $\epsilon = 80$, respectively. The nucleophilic attack results in a penta-coordinated intermediate (Inter, Figure 5.1), which is calculated to be located in a shallow minimum, ca 1.5 kcal/mol lower than TS1. The P–O$_\lambda$ bond breaks with a barrier of only ca 2 kcal/mol relative to the
intermediate (ca 12 kcal/mol relative to the reactant species, see TS2 in Figure 5.1). It turns out that, simultaneously with the P-O₅ bond cleavage, the proton of the bridging hydroxide is transferred to the Asp301 residue, resulting in a phosphate anion bridging the two zinc ions in the enzyme-product complex (Prod, Figure 5.1).

In this step, the CPCM solvation effects are more significant, because the leaving group departs as an anion. The combined exothermicity of the two steps is calculated to be 4.6 kcal/mol in the cluster model, and 12.6 and 15.3 kcal/mol when ε = 4 and ε = 80 are applied, respectively (see Figure 5.2). It is interesting to monitor how the bond distances of the bridging hydroxide to the two zinc ions change during the reaction. From two symmetric bonds of 1.99 and 2.00 Å in the reactant species, the distances increase to 2.22 and 2.15 Å to β and α metals, respectively, in the penta-coordinated intermediate structure. In the product species, the bond to the β-Zn is broken completely (3.44 Å) while the bond to the α-Zn is maintained (2.04Å). It is worth noting that this calculated structural feature in Prod is in excellent agreement with the crystal structure of PTE-product complex recently solved by Raushel et al (see Paper III).

Another interesting parameter is the distance between the phosphoryl oxygen and the β-Zn. From a quite weak binding of 2.24 Å in the enzyme-substrate complex, the distance is decreased to 2.04 Å in the penta-coordinated intermediate, and 1.99 Å in the enzyme-product complex. As concluded in Section 4.4.3, these bond lengths indicate that the role of the β-Zn ion is to stabilize the intermediate and product structures, thereby lowering the barriers for the nucleophilic attack and the P-O₅ bond cleavage steps. This explains the increased nucleophilicity of the bridging hydroxide, a property that has been questioned[51].

Figure 5.2 Potential energy profile for the PTE reaction with dimethyl 4-nitrophenyl phosphate as a substrate. (Cluster) B3LYP/6-311+G(2d,2p) energies with zero-point energies included. (CPCM) CPCM solvation effects added with different dielectric constants (ε = 4 or 80).
5. Applications

5.1.3 Base mechanism for dimethyl 4-nitrophenyl phosphate substrate

To assess the feasibility of the base mechanism proposed by Ollis et al. (Scheme 4.2)[54], we have performed quantum chemical calculations (Paper III) using the same Zn/Zn-PTE active site model as in the calculations of the nucleophile mechanism (shown in Figure 3.3). Dimethyl 4-nitrophenyl phosphate is used as the substrate in this investigation. The water molecule, postulated to be the nucleophile in the base mechanism, was added to the model. The model thus consists of 85 atoms and has a total charge of +1.

Figure 5.3 Optimized geometries of the Zn/Zn-PTE active site model with dimethyl 4-nitrophenyl phosphate and a water molecule bound (React), and the transition state for the water attack on the phosphorus (TS).

In the reactant species (Figure 5.3, React), the phosphoryl oxygen of the substrate binds to the more solvent-exposed β-Zn site and the leaving group is located at the position opposite to the bridging hydroxide. According to the base proposal, the nucleophilic water molecule is bound to the α-Zn. In the calculations, we first placed the water molecule there, but we could not locate a stable conformation with the water bound to the α-Zn. Instead, it dissociates from the Zn during the optimization and forms a hydrogen bond to the bridging hydroxide (see Figure 5.3, React). From these calculations, it can thus be concluded that a hexa-coordinated α-Zn is quite unlikely.

We have furthermore optimized the transition state for the nucleophilic attack with the bridging hydroxide acting as a base. The optimized structure (TS) is displayed in Figure 5.3. It can be found that, simultaneously with formation of the P-O bond, one proton of the water molecule is transferred to the bridging oxygen. This indicates that the O$_\mu$H$^-$ serves as the general base to activate the attacking water. This attack results in the penta-coordinated intermediate. The barrier and reaction energy for this step are calculated to be 17.3 and 13.3 kcal/mol, respectively. The barrier for the base mechanism is thus significantly higher than the barrier where the bridging hydroxide is the nucleophile (11.7 kcal/mol, see Section 5.1.2 and Paper I). This finding thus provides additional evidence that the bridging hydroxide acts as a nucleophile and not as a general base in the Zn/Zn-PTE reaction.
5. Applications

5.1.4 Trimethyl phosphate substrate

Trimethyl phosphate is not the substrate for PTE experimentally. To provide a deeper insight into the PTE catalysis, we used trimethyl phosphate as a substrate and optimized the stationary points along the reaction pathway. In this study, only nucleophile mechanism was considered.

As for the hydrolysis of dimethyl 4-nitrophenyl phosphate substrate, the first step for trimethyl phosphate substrate is the nucleophilic attack on the phosphorus center by the $\text{O}_\mu\text{H}^-$. The transition state (TS1 in Figure 5.4) has been calculated to be a first-order saddle point with an imaginary frequency of 55.6$\text{ i cm}^{-1}$. At TS1, the distances of the P to the $\text{O}_\mu$ and the $\text{O}_\text{L}$ are 2.07 and 1.66 Å, respectively, to be compared to the values of 2.24 and 1.68 Å in the corresponding transition state of dimethyl 4-nitrophenyl phosphate hydrolysis (see TS1 of Figure 5.1). The barrier height of this step is predicted to be 10.9 kcal/mol with the solvation ($\varepsilon = 4$) included (see Figure 5.5), indicating that the nucleophilic attack is a feasible step. It is interesting to note that this barrier is quite close to that of dimethyl 4-nitrophenyl phosphate hydrolysis (10.8 kcal/mol, see Figure 5.2). This can be rationalized by the reaction nature that nucleophilic attack basically depends on the nucleophilicity of the $\text{O}_\mu\text{H}^-$ and the electrophilicity of the phosphorus center, but not the character of the leaving group.

The second step of the P-$\text{O}_\text{L}$ bond cleavage is quite different from the dimethyl 4-nitrophenyl phosphate system. The optimized transition state (TS2 in Figure 5.4) is very late with a large P-$\text{O}_\text{L}$ distance of 2.76 Å. The methoxide is a bad leaving group due to its high $pK_a$. As a result, the barrier of this step is calculated to be very large (40.2 kcal/mol, $\varepsilon = 4$, Figure 5.5). This means that the P-$\text{O}_\text{L}$ bond cleavage could not take place for trimethyl phosphate substrate. On the basis of these calculations, it can be concluded that the high $pK_a$ of the leaving group increases the barrier of the P-$\text{O}_\text{L}$ bond cleavage, thus blocking the hydrolysis of trimethyl phosphate.
5. Applications

21.9
37.0
11.1
11.8
31.4
0.0
10.7
11.7
48.7
24.4
36.1
40.2

Figure 5.5 Potential energy profile for the PTE reaction with trimethyl phosphate as a substrate. (Cluster) B3LYP/6-311+G(2d,2p) energies with zero-point energies included. (CPCM) CPCM solvation effects added with different dielectric constants (ε = 4 or 80).

5.1.5 Conclusions

Our calculations provide strong support to an associative step-wise mechanism for the hydrolysis of phosphotriesters by PTE, where the bridging hydroxide functions as the direct nucleophile and a penta-coordinated intermediate is formed. No activation (protonation) of the leaving group was found necessary. As commented in Section 4.1, it can be concluded that the nucleophilicity of the bridging hydroxide is sufficient to be utilized in the hydrolysis reaction, a feature that is of importance for a number of other di-zinc enzymes[73]. We have, furthermore, demonstrated that the bridging hydroxide acts as a direct nucleophile and not as a general base in the Zn/Zn-PTE reaction. Also, our calculations reveal that the high pKa of the methoxide largely increases the barrier of the P-O_l bond cleavage step, making the hydrolysis of trimethyl phosphate impossible without further assistance.

5.2 Aminopeptidase from Aeromonas proteolytica (AAP) (Paper IV)

Metalloaminopeptidases represent a family of enzymes that use one or two metal ions to specifically cleave the N-terminal amino acid residues of polypeptides and proteins[74]. They play fundamental roles in different biochemical events, such as protein maturation and degradation, tissue repair, and cell-cycle control. Aminopeptidase from Aeromonas Proteolytica (AAP) is a monomeric enzyme (30 kDa) containing two Zn²⁺ ions that are essential for its enzymatic activity[75]. AAP exhibits a substrate preference for the hydrophobic N-terminal amino acid residues[76] with a turnover number (k_cat) close to 90 s⁻¹ for the L-leucine-p-nitroanilide substrate at room-temperature[77].
5.2.1 Structure and mechanism

The X-ray crystal structure of AAP (PDB entry 1RTQ) reveals that the active site contains a binuclear zinc center that is bridged by an aspartate (Asp117) and an oxygen species (see Figure 3.2), which from the high resolution X-ray structure has been established as a hydroxide \((\cdot\cdot\cdotH\cdot\cdot\cdot)\)\(^7\). In addition, the two zinc ions are coordinated to the protein via the side chains of His97, Asp179, Glu152, and His256. Glu151, a second-shell residue, is situated very close to the di-zinc core, forming a hydrogen bond to the bridging hydroxide.

Based on the crystallographic studies and other kinetic, spectroscopic, and mutational studies, the following reaction mechanism has emerged for AAP. Once the substrate docks onto the active site, the carbonyl group of the peptide bond is proposed to coordinate to Zn1\(^7\). Next, N-terminal amino group binds to Zn2, which induces the bridging hydroxide to become terminal, bound only to Zn1\(^7\). The resulting Zn1-OH then performs a nucleophilic attack at the carbonyl group leading to the formation of a tetrahedral gem-diolate intermediate. Next, a proton is transferred from the second-shell Glu151 to the nitrogen, resulting in the breakdown of the tetrahedral intermediate\(^8\). The active site is finally regenerated by releasing the products and binding a new bridging ligand.

Quantum mechanical/molecular mechanical (QM/MM) calculations have been used to study the reaction mechanism of AAP\(^5\). This study gave general support to the mechanism outlined above. In particular, these calculations confirmed the notion that the bridging hydroxide dissociates and becomes terminal before it can attack the substrate carbonyl. The QM part of the calculations was, however, chosen to be quite large, which only allowed the system to be treated at the semi-empirical AM1 level of theory. This drawback permitted thus only qualitative description of the energetics of the mechanism, since the obtained barriers were very high.

5.2.2 Active site model

A model of the active site of AAP (shown in Figure 3.2) was constructed on the basis of the crystal structure of the native enzyme (PDB entry 1RTQ)\(^7\). As a substrate, we chose a simple model of a di-alanine peptide. The C-terminal of the dipeptide was blocked with a methyl group to avoid artificial interactions of the carboxyl group with the active site. On the basis of previous results of experimental and theoretical work\(^8\), some assumptions have been made about the protonation states of the various groups. The amino terminal of the substrate is assumed to be neutral and the bridging ligand is a hydroxide. Glu151 was protonated in the starting structure, consistently with its proposed role of activating the bridging water by abstracting a proton\(^1\). The total number of atoms in the model is 82 and the total charge is zero.

The overall geometric parameters of reactant obtained from the geometry
optimization (see Figure 3.2) agree quite well with the X-ray structure. It is interesting to note that, in this reactant structure, the amino group of the substrate binds to the Zn2 ion (Zn2-N distance is 2.37 Å), while there is no interaction between Zn1 and the peptide bond oxygen\(^1\). A possible reason for this could be that Glu152 binds to Zn1 in a bidentate fashion, making coordination of the carbonyl group to the same zinc less favorable as compared to the coordination of the amino group to Zn2, to which Asp179 is bound in a monodentate fashion. Similar results were obtained in the AM1/MM study\(^{[85]}\). As discussed in Section 4.4.1, this indicates that the Zn2 site plays a role of assisting in orientating the peptide bond toward the nucleophile (O\(_\mu\)H).

5.2.3 Nucleophilic attack

From React, we find that the hydroxide attack happens directly from bridging position, and that the dissociation of the bridging hydroxide from Zn1 takes place simultaneously at the same transition state. The optimized transition state for this reaction step (TS1) and the resulting intermediate (Int1a) are displayed in Figure 5.6. The barrier of this step is calculated to be 14.6 kcal/mol with the solvation effects included, and the intermediate lies 13.4 kcal/mol higher than the reactant.

Some important chemical information can be reflected by the geometrical changes that take place in going from React to Int1a, via TS1 (see Figures 3.2 and 5.6). The peptide carbonyl double bond is elongated from 1.23 to 1.35 Å (1.30 Å at TS1), and the developing charge at the oxygen leads to its coordination to the Zn1 ion (1.98 Å in TS1 and 1.92 Å in Int1a). This demonstrates that Zn1 provides catalytic power by stabilizing the developing negative charge of the oxygen atom of the tetrahedral intermediate, as pointed out in Section 4.4.3. The bridging O\(_\mu\) becomes terminal with distances of 2.07 Å to Zn2 and 2.74 Å to Zn1 in Int1a. The resulting gem-diolate moiety can thus be considered as a bridging ligand between the two Zn ions. The distance of Zn2 to the N-terminal nitrogen is also slightly shortened in the tetrahedral intermediate compared to React (2.24 compared to 2.37 Å), which indicates that the Zn2 ion, in addition to binding and orienting the substrate, might also contribute to the catalysis.

\(^1\) In the following transition state and intermediate state of the nucleophilic attack step, Zn1 center will bind the carbonyl oxygen and stabilize the developing charge.
5.2.4 Proton transfer

The second step in the hydrolysis reaction is the proton transfer from the Glu151 to the nitrogen of the peptide bond (N_P). To be able to do this, Glu151 has to rotate away from Asp179, and thus break its hydrogen bond to that residue. This rotation results in a new intermediate, denoted Int1b, which lies 4.1 kcal/mol higher than Int1a (structure displayed in Figure 5.6).

The transition state for the subsequent proton transfer (TS2) and the resulting intermediate (Int2) were optimized and are shown in Figure 5.7. This proton transfer turns out to be the rate-limiting step for the hydrolysis, with a barrier of 18.2 kcal/mol relative to the reactant species of Figure 3.2, i.e. only less than 1 kcal/mol higher than Int1b. The resulting Int2 intermediate is calculated to be 15.0 kcal/mol higher than the reactant. At Int2, the distance between the proton and N_P is 1.08 Å. As the nitrogen gets protonated, the C-N bond for the peptide is elongated from 1.50 to 1.58 Å (1.53 Å at TS2). Also, the hydrogen bond between Glu151 and the O_H becomes stronger as Glu151 becomes anionic.
5.2.5 C-N bond cleavage

At Int2, the peptide bond is in principle ready to cleave. However, it turns out that another proton transfer step must take place before the C-N bond can dissociate. Every attempt to optimize a transition state for the C-N bond cleavage from Int2 resulted in a transition state for a proton transfer from the O$_\mu$H' to Glu151. The
optimized geometries of this transition state (TS3) and the resulting intermediate (Int3) are shown in Figure 5.7. Int3 is only slightly more stable than Int2 (0.6 kcal/mol). The barrier is calculated to be slightly negative, i.e. the energy of TS3 is slightly lower than the energy for Int2 (0.6 kcal/mol). This is of course an error of the methods and models employed. When the real barrier is very low, a small underestimation in the calculated barrier can yield artificially slightly negative barriers. The firm conclusion, however, is that this step is very fast.

From Int3 a transition state for the C-N bond cleavage could be located (TS4). Also for this step, the barrier is calculated to be slightly negative (ca 0.6 kcal/mol), which again is an error of the methods, but which shows that the step is very fast. The C-N bond cleavage leads to a neutral amino group which dissociates and a deprotonated carboxylate moiety that is bridging the two zinc ions (see Figure 5.7, Prod). The energy of this species is 4.6 kcal/mol higher than that for the reactant species of Figure 5.7.

It should here be mentioned that several attempts were made to find a concerted transition state, i.e. proton transfer from Glu151 coupled with the C-N bond cleavage, but without success.

5.2.5 Conclusions

In this study, a reaction mechanism for AAP can be obtained (presented in Scheme 5.1) and the corresponding PES is presented in Figure 5.8. With help of the calculations, the following mechanistic features could be established. First, the substrate is found to bind to the di-zinc cluster through the amino group and not the peptide carbonyl. The Zn2 ion has thus an important role in binding the substrate and orienting the peptide bond toward the nucleophile. The calculations furthermore demonstrate that the bridging hydroxide is capable of performing a nucleophilic attack at the peptide carbonyl without the need to become terminal. This conclusion is of importance for the other hydrolytic di-zinc enzymes\cite{73}, and has been remarked in Section 4.1. The main catalytic power originates from the Zn1 ion, which stabilizes the anionic tetrahedral intermediate thereby lowering the barrier for the nucleophilic attack, as pointed out in Section 4.4.3. The rate-limiting step is found to be the proton transfer from Glu151 to the nitrogen of the peptide bond, with an accumulated barrier of 18.2 kcal/mol. This is consistent with the experimental finding that the rate-limiting step involves a proton transfer\cite{83,86}.
5. Applications

**Scheme 5.1** Proposed reaction mechanism for AAP based on the present calculations.

**Figure 5.8** Calculated potential energy profile for peptide hydrolysis by AAP. (Cluster) B3LYP/6-311+G(2d,2p) energies with zero-point energies included. (PCM) IEF-PCM solvation effects added with dielectric constant $\varepsilon = 4$. 
5.3 Glyoxalase II (GlxII) (Paper V)

The glyoxalase system, composed of two enzymes (glyoxalase I and II), catalyzes the conversion of toxic 2-oxoaldehydes to the corresponding 2-hydroxycarboxylic acids using glutathione (GSH) as a cofactor\cite{87}. In particular, the system can detoxify methylglyoxal, a byproduct of carbohydrate and lipid metabolism\cite{88,89}, which can produce toxic effects by reacting with RNA, DNA, and proteins\cite{88,90}. Glyoxalase I converts thiohemiacetal (formed from methylglyoxal and glutathione) to S-D-lactoylglutathione, while Glyoxalase II (GlxII) takes the latter as a substrate and catalyzes its hydrolysis to yield D-lactic acid and GSH (Scheme 5.2). S-D-lactoylglutathione is also known to exhibit cytotoxic effects through the inhibition of DNA synthesis\cite{88}. The glyoxalase system is thus thought to play an indispensable role in chemical detoxification\cite{88,91}.

While GlxI is a mononuclear zinc enzyme\cite{92}, GlxII utilizes a binuclear metallosite to perform its reaction. It has a turnover number \(k_{\text{cat}}\) of \(2.8 \times 10^2\) s\(^{-1}\) and a \(k_{\text{cat}}/K_m\) value of \(8.8 \times 10^5\) M\(^{-1}\)s\(^{-1}\) for S-D-lactoylglutathione substrate\cite{93}. GlxII generally requires two divalent zinc ions for catalytic activity\cite{94,95}. However, this is not an absolute requirement. Glyoxalase II from *Arabidopsis thaliana*\cite{96,97} and *Salmonella typhimurium*\cite{98}, for example, show metal-binding flexibility in that a variety of bimetallic sites, with iron, manganese, or zinc embedded, have been characterized.

![Scheme 5.2 Reactions catalyzed by the Glyoxalase system.](image)

5.3.1 Structure and mechanism

Several X-ray crystal structures of GlxII have been deposited\cite{99-101}, including human GlxII (PDB entry 1QH3) and its complex with a glutathione thioester substrate analogue (PDB entry 1QH5)\cite{99}. The active site of GlxII contains a binuclear zinc center that is bridged by one oxygen of the Asp134 and a hydroxide (O\(_{\mu}\)H\(^-\))\cite{99}. The two zinc ions are bound to the protein via the side chains of His54, His56, His110, Asp58, His59, and His173. The Asp58 residue forms a hydrogen bond to the bridging hydroxide, suggesting that it plays a role in orienting the hydroxide for attacking on the substrate\cite{100}, as mentioned in Section 4.5.

Based on extensive experimental work, the following mechanism (see Scheme 5.3) has been proposed for the GlxII reaction\cite{99,100}. The S-D-lactoylglutathione substrate docks first onto the active site, and the carbonyl oxygen of lactoyl group (O\(_C\)) is proposed to coordinate to Zn2 ion\cite{99}. The bridging hydroxide (O\(_{\mu}\)H) then performs a nucleophilic attack at the carbonyl carbon of the lactoyl group (C\(_C\)) leading to a tetrahedral gem-diolate intermediate\cite{99,100}. Next, the C\(_C\)-S bond of the substrate breaks with help of the Zn1 ion\cite{100}, resulting in a D-lactic acid and an unprotonated
5. Applications

Glutathione (GS\(^{-}\)). To complete the reaction, the GS\(^{-}\) group is protonated by a water molecule and the products are released from the active site to regenerate the enzyme.

**Scheme 5.3** Proposed reaction mechanism for GlxII.

### 5.3.2 Active site model

A model of the active site of GlxII (displayed in **Figure 5.9**) was constructed on the basis of the crystal structure of the wild-type *human* GlxII in complex with a slow substrate, S-(N-hydroxy-N-bromophenylcarbamoyl) glutathione (PDB entry 1QH5)\(^{[99]}\). The model contains the two zinc ions along with their first-shell ligands and the bridging Asp134 and hydroxide. As a model substrate, we chose a small thioester formed from lactic acid and methylthiol (see **Figure 5.9**). The total number of atoms in the model is 78 and the total charge is +1.

The optimized structure of the GlxII active site with the model substrate bound, is displayed in **Figure 5.9**. The overall geometric parameters obtained from the geometry optimization agree quite well with the X-ray structure. It is interesting to note that the substrate does not coordinate to any of the zinc ions in the Michaelis complex. Instead, the hydroxyl group forms a hydrogen bond to the Asp58 residue. This hydrogen bond could be artificial and caused by the relative smallness of the...
model. In the real enzyme, one can envision that the extended substrate binds to the enzyme through other groups not present in the models of the active site or the substrate, such as the glycine and cysteine parts of the substrate. However, these interactions would be constant and do not change during the reactions. They contribute thus only to the binding, which makes the use of the current model justifiable for the purpose of studying the chemical steps of the reaction.

![Figure 5.9](image)

*Figure 5.9* Optimized structure of the GlxII active site with the model substrate bound. For clarity, unimportant hydrogen atoms are omitted, which is also applied to other figures below. Arrows indicate the atoms that are fixed to their X-ray positions. Inserted is the model substrate used in the calculations.

### 5.3.3 Nucleophilic attack.

The first step of the proposed mechanism is the nucleophilic attack of the bridging hydroxide on the carbonyl carbon (C₆) of substrate[99,100]. We have located the transition state (called **TS1**, see *Figure 5.10*) for this nucleophilic attack. We see that it takes place directly from the bridging position. The energy barrier is calculated to be 19.2 kcal/mol upon addition of surrounding solvation in the form of CPCM. The resulting tetrahedral intermediate (**Inter1**) is calculated to be 19.1 kcal/mol higher than **React**, i.e. only 0.1 kcal/mol lower than **TS1**.
Figure 5.10 Optimized structures of transition states and intermediates along the reaction pathway of GlxII.

As commented in Section 4.4.3, the distance changes of the C\textsubscript{C} to the O\textsubscript{μ} and carbonyl oxygen (O\textsubscript{C}) demonstrate that Zn2 provides catalytic power by stabilizing the emerging negative charge at the O\textsubscript{C} atom of the tetrahedral intermediate. Another interesting observation here is that since both O\textsubscript{μ} and Asp134 remain bridging, Zn2
becomes hexa-coordinated. This is quite unusual for zinc enzymes, and is probably related to the fact that Asp134 bridges the two zinc ions with one oxygen atom. In other dinuclear zinc enzymes, such as PTE, AAP, and DHO\[^{102}\], the carboxylate group bridges by one oxygen to each zinc ion. Finally, we note that the sulfur atom of the substrate starts to associate with Zn1. The Zn1-S distance is 2.65 Å at \textbf{Inter1}. This interaction will become important for the S-C bond cleavage step (see below).

\subsection*{5.3.4 C-S Bond cleavage}

Next step after the formation of \textbf{Inter1} is the C-C-S bond cleavage\[^{99,100}\]. From the calculations, it turns out that a proton transfer step must take place before the C-S bond can dissociate. Every attempt to locate a transition state for the C-S bond cleavage from \textbf{Inter1} led to a transition state (called \textbf{TS2}) for a proton transfer from the O\(_{\mu}\)H to Asp58. The barrier and reaction energy for this proton transfer step are calculated to be very small compared to \textbf{Inter1}, +0.3 and +0.2 kcal/mol, respectively.

At \textbf{Inter2}, the C-C-S bond is ready to break. We have located a transition state for the C-S bond cleavage from \textbf{Inter2} (\textbf{TS3}, see Figure 5.10). Although the barrier is calculated to be only 0.1 kcal/mol higher than \textbf{Inter2}, it is a distinct stationary point with an imaginary frequency of 79i cm\(^{-1}\). The resulting enzyme-product complex (\textbf{Prod}), is 7.5 kcal/mol lower than \textbf{Inter2} (i.e. 11.8 kcal/mol higher than \textbf{React}). The C-S bond cleavage results in a thiolate (CH\(_3\)S–) that coordinates to Zn1 and a lactate that coordinates to Zn2 in a bidentate fashion (see \textbf{Prod}). The charge at the sulfur leads to a stronger coordination to the Zn1 ion (2.50 Å in \textbf{TS3} and 2.38 Å in \textbf{Prod}). As mentioned in Section 4.4.3, this result indicates that the Zn1 ion plays a role in stabilizing the developing charge at the sulfur, thereby lowering the barrier for the C-C-S bond cleavage step. With the S-Zn1 bond formation, O\(_{\mu}\) loses coordination to Zn1 (O\(_{\mu}\)-Zn1 = 2.80 Å in \textbf{Prod}), which results in a maintained penta-coordination of Zn1 and hexa-coordination of Zn2.

\subsection*{5.3.5 Product release and active site regeneration}

At \textbf{Prod}, both products (thiolate and lactate) are negatively charged and strongly bound to their respective zinc ions (see Figure 5.10). To regenerate the active site for another round of catalysis, the products have to be released and a new water molecule has to enter and form a bridging hydroxide. To obtain exact energies of this process is very complicated and requires much larger models of the active site and also very accurate bulk solvent representation. However, from our calculations, one can make a rough estimation of the energetics involved in the ligand exchange process by using the following procedure.

We first note that the overall reaction of GlxII:

substrate + water → methylthiol + lactic acid
is calculated to be almost thermoneutral (exothermic by 1.5 kcal/mol in the gas phase and endothermic by 0.4 kcal/mol when solvation is considered using $\varepsilon = 80$). Since the enzyme does not change this energy, one full enzyme cycle should also be endothermic by this amount (for a general introduction about the energetics of the enzyme catalysis see Section 3.1). Thus, one can estimate that the steps involved in the product release, active site regeneration, and the binding of a new substrate to the active site, together are exothermic by ca 11.4 kcal/mol, which is the energy of Prod added to the overall reaction energy (11.8 – 0.4 kcal/mol).

Furthermore, to estimate approximately the energies involved in the individual steps, we first let a water molecule bind in the second coordination shell of Prod and optimized the geometry of the complex (called Prod-wat, Figure 5.11). Subsequently, we let this water molecule change places with either of the two products (methylthiol and lactic acid), which now bind in their protonated forms in the second coordination shell. The optimized structures of these complexes are called Comp1 and Comp2, respectively. Both these structures have slightly higher energies than Prod-wat, Comp1 by 0.3 kcal/mol and Comp2 by 4.8 kcal/mol.

In the case of Comp1 we notice that in the geometry optimization the lactate product takes a proton from the protonated Asp58 rather than from the newly-introduced water molecule (see Figure 5.11). When the lactic acid is removed from the model and the geometry is re-optimized (Comp1'), we found that the Asp58 picked up one proton from the water molecule and the resulting hydroxide became bridging between the two zinc ions. To regenerate the active site, the thiolate might take the proton of Asp58 and get released as a neutral molecule.

Alternatively, when the methylthiol product is removed from the Comp2 and the structure was re-optimized, something very interesting is observed. We found that the Zn1-stabilized hydroxide ($O_{Zn1}H^-$) without a barrier attacked the $C_C$ atom of the lactate and formed a new tetrahedral intermediate (see Comp2'). Simultaneously with the $O_{Zn1}-C_C$ bond formation, a proton is delivered back from the Asp58 residue to $O_\mu$. From Comp2', it is possible that the $O_\mu-C_C$ bond breaks and the active site is regenerated this way. We therefore optimized the TS for the $O_\mu-C_C$ bond cleavage and the barrier was calculated to be only 2.0 kcal/mol relative to Comp2', indicating that this step is feasible. This step, resulting in a free lactic acid and a regenerated active site, is calculated to be exothermic by 18.1 kcal/mol. This outlined scenario for the active site regeneration is quite new. However, it should be stressed that it is very speculative, since it is based on quite crude estimations. These steps cannot be modeled accurately using the models and methods employed in the current study. More experimental and theoretical work will be required to make a more solid statement about the exact mechanism of active site regeneration.
Figure 5.11 Optimized structures of the complexes in the ligand exchange processes.
5. Applications

![Reaction Energy Profile](image)

**Figure 5.12** Calculated potential energy profile for GlxII reaction. (Cluster) B3LYP/6-311+G(2d,2p) energies with zero-point corrections included. Solvation effects are added with dielectric constant $\varepsilon = 4$.

### 5.3.6 Conclusions

The calculations performed on the GlxII active site model give general support to the previously proposed mechanism of **Scheme 5.3**\(^{99,100}\) and provide a more detailed picture of the various steps of the reaction. The calculated PES is shown in **Figure 5.12**. To summarize, the following mechanistic characteristics can be concluded from the calculations:

Similarly to previous calculations on other dinuclear zinc enzymes, such as PTE (**Section 5.1**), AAP (**Section 5.2**), and Dihydroorotase (DHO)\(^{102}\), the calculations demonstrate that the bridging hydroxide can indeed act as a nucleophile, attacking the carbonyl carbon of lactoyl group directly from its bridging position.

By studying how some important geometric parameters change during the reaction we were able to analyze the roles of the two zinc ions. Zn2 is demonstrated to stabilize the charge of the tetrahedral intermediate, lowering thereby the barrier of the nucleophilic attack. Zn1, on the other hand, stabilizes the emerging charge at the thiolate, facilitating the S-C bond cleavage step. These functions of zinc ions are crucial for the hydrolysis catalyzed by the binuclear zinc hydrolases and have also been commented on in **Section 4.4.3**.

The nucleophilic attack has a calculated barrier of 19.2 kcal/mol, while the later S-C bond dissociation step has a barrier of 19.4 kcal/mol. The two barriers are thus
too close to make a solid conclusion from the calculations regarding which one is the rate-limiting. However, in at least three other binuclear zinc hydrolases (PTE, AAP, and DHO), calculations have shown that the initial nucleophilic attack by the bridging hydroxide is not the rate-limiting step. It is possible that this is a general feature in binuclear zinc hydrolases. If this is the case, the results obtained for the GlxII reaction are not in contradiction with this.

Finally, the present study has also considered the steps involved in the product release and active site regeneration. A new possible mechanism was outlined, in which the thiol product is released and replaced by a Zn-bound hydroxide. This hydroxide then attacks the lactate product, bound to the other zinc ion, to form a tetrahedral intermediate (Comp2′), which then collapses to yield the lactic acid and a regenerated active site with a bridging hydroxide (see **Figure 5.11**). However, due to the smallness of the active site model employed, the energy considerations are very crude and the conclusions of this part of the study are only speculative.

### 5.4 Alkaline Phosphatase (AP) (Paper VI)

Alkaline phosphatase (AP) catalyzes the hydrolysis of a broad range of phosphate monoester substrates to produce inorganic phosphate and alcohol (or phenol) under both acidic and alkaline conditions\[^{103-105}\]. It is an almost perfect, i.e., nearly diffusion-limited, enzyme\[^{106,107}\], with a $k_{cat}/K_m$ value of $(3.3±0.5)\times10^7$ M$^{-1}$s$^{-1}$ for p-nitrophenyl phosphate substrate at room temperature\[^{108}\]. AP also displays some catalytic promiscuity with a low level of phosphodiesterase activity\[^{109,110}\]. It has been an interesting system for characterization of the transition states during the enzyme-catalyzed hydrolysis of phosphate esters and investigation of the catalytic power of binuclear zinc hydrolases\[^{103,111-118}\].

#### 5.4.1 Structure and mechanism

Recently, a new crystallization strategy was used to obtain the crystals of AP with and without the complexed inorganic phosphate (Pi) at high resolution (1.75 Å)\[^{119}\], where the PDB codes are 1ED8 and 1ED9, respectively. As seen in **Figures 5.13** and **5.14**, two zinc ions and one magnesium ion are contained in the active site of AP, but the two zinc ions are not connected by a bridging hydroxide as the three other enzymes studied in this thesis. The two zins are bound to the protein via the side chains of His331, His412, Asp327, Asp369, and His370. The carboxylate group of Asp51 acts as a bridge to connect Zn1 with Mg$^{2+}$. The Mg ion binds with octahedral geometry and, besides Asp51, it is coordinated by Glu322, Thr155, and three water molecules. In the crystal structure of AP in complex with inorganic phosphate, Pi sits in the pocket between two zinc ions and is fastened by the Arg166 via two hydrogen bonds\[^{109}\]. Mutation of this arginine residue does not inactivate the enzyme\[^{109,120}\] but demonstrates that the Arg166 is possibly accountable for the preference of AP for phosphate monoesters\[^{109}\]. The Ser102 residue, which is reported to have a pK$_a$ of
in the free enzyme, appears to act as a crucial nucleophile to attack the phosphorus center during the reaction. One of the water molecules is possibly deprotonated to a hydroxide ion ($\text{O}_\text{MgH}^-$, see $\text{R}_\text{me}$ and $\text{R}_\text{pnp}$ in Figures 5.13 and 5.14, respectively) by the solution and is expected to work as a general base/acid for regulating the protonation state of Ser102.

Based on these X-ray crystal structures, a three-metal reaction mechanism was proposed by Stec et al., that suggests two in-line displacements occur at phosphorus center (see Scheme 5.4). The first in-line attack comes from the Ser102 alkoxide, which is activated by the hydroxide ($\text{O}_\text{MgH}^-$) stabilized by the Mg ion. This drives out a negatively charged leaving group (alkoxide or phenolate) possibly stabilized by Zn2 ion. This displacement step results in a covalent phosphoseryl intermediate (PSI) and is thus called “phosphoryl transfer”. Then, a
solvent water molecule comes in and protonates the leaving group, leading to free alcohol (or phenol) product and a Zn2-bound hydroxide (OZn2H). Next, the resulting OZn2H performs the second in-line attack in order to produce the enzyme-Pi product complex\cite{119}, simultaneously regenerating the Ser102 and OmgH\cite{103,119}. This step is named by “hydrolysis of PSI”. Apparently, this Ping-Pong mechanism proceeds with retention of the configuration at phosphorus center, which has been confirmed by isotope-labeling experiments\cite{105}.

A necessary attempt to further understand enzymatic catalysis is the characterization of the transition state during the enzyme reaction. A number of studies using linear free energy relationships (LFERs, see the brief introduction in Section 4.3) have been performed with respect to AP reaction, but provide quite different information about the nature of the transition states\cite{108,126-129}. For example, the AP-catalyzed hydrolysis of aryl O-phosphorothioates has been observed to have a steep leaving group dependence (the Bronsted value β = -0.77)\cite{126}, indicating that the transition state is largely dissociative\cite{111,126}. In contrast, the studies using aryl phosphate substrates showed that the turnover number (kcat) was almost independent of the pKa of the leaving group and nearly constant\cite{108,127-129}. For alkyl phosphates, the earlier studies found little dependence upon the nature of the leaving group\cite{127,129}, but a recent study employing a more sensitive 32P-based assay argued to the contrary, that the dependence of kcat/Km on the pKa of the leaving group followed a steep Bronsted correlation with β = -0.85\cite{108}. Furthermore, the LFERs experiments listed above were not able to distinguish between concerted and step-wise reaction mechanisms.

5.4.2 Active site model

A model of the active site of AP was built on the basis of the high resolution crystal structure (PDB code 1ED8)\cite{119}. The model contains the two zinc ions, the magnesium ion, His331, His412, Asp327, Asp369, His370, Asp51, Glu322, Thr155, Arg166, Ser102, two water molecules ligated to Mg, and a Mg-stabilized OH- (OmgH). The negatively charged Asp101, a second-shell residue that forms strong hydrogen bonds to the positively charged Arg166, is also included in the model. A scrutiny of the hydrolysis of only one substrate is, however, not sufficient to give a panorama of AP reaction, since the enzymatic hydrolysis of phosphate esters has been shown to be rather complicated, where even a small alteration of the acidity of the leaving group may change the transition state nature\cite{130,131}. Therefore, in this investigation, we take two substrates, methyl and p-nitrophenyl phosphates, which are considered as the representatives of alkyl and aryl phosphate substrates, respectively. The total charge of the model is -1 and the total numbers of atoms are 110 and 119 for the methyl and p-nitrophenyl phosphate systems, respectively.
Figure 5.13 Optimized structures of the stationary points along the phosphoryl transfer pathway in the methyl phosphate hydrolysis and the corresponding More O’Ferral Jencks (MFJ) plot. For clarity, some hydrogen atoms are omitted, which is also applied to other figures below. Stars indicate the atoms that are fixed to their X-ray positions. In the MFJ plot, the key distances of the phosphor to the leaving group oxygen (P-O_{lg}) and the nucleophile oxygen (P-O_{nuc}) are given in the square brackets, and the relative energies (kcal/mol) are provided in the parentheses.

The optimized structures of the AP active site with methyl and p-nitrophenyl phosphate substrates bound, corresponding to the Michaelis complexes and respectively referred to as R_{me} and R_{pnp} hereafter, are displayed in Figures 5.13 and 5.14, respectively. The overall geometric parameters obtained from the geometry optimizations agree quite well with the X-ray structure. For example, the Zn1-Zn2 distances are calculated to be 4.06 and 4.00 Å for R_{me} and R_{pnp}, respectively, to be compared to the crystallographic distance of 3.98 Å\[119\]. One of the phosphoryl oxygens, denoted by O_{μ}, acts as a bridge to connect the two zinc ions.
5.4.3 Methyl phosphate substrate

In the phosphoryl transfer step, the nucleophile (Ser102 alkoxide) is supposed to be activated by the hydroxide bound to Mg\(^{113,117,119}\). However, the origin of this crucial general base (O\(_{\text{MgH}}\)) is still an open question. The most likely possibility is that the O\(_{\text{MgH}}\) derives from the deprotonation of a water molecule ligated to Mg by the hydroxide in solution. A similar example has been demonstrated in phosphodiester hydrolysis catalyzed by DNA Polymerase I\(^{123}\).

From the AP-methyl phosphate complex (see \(R_{\text{me}}\) in Figure 5.13), we have optimized the transition state (called \(T\_{\text{S,me}}\)) for the nucleophilic attack on the phosphor by the Ser102 oxygen (O\(_{\text{nuC}}\)). The optimized structures of the \(T\_{\text{S,me}}\) and the resulting phosphoseryl intermediate (\(P\_{\text{S,me}}\)) are also shown in Figure 5.13. The nature of \(T\_{\text{S,me}}\) has been confirmed to be a first-order saddle point with an imaginary frequency of 116\(i\) cm\(^{-1}\), which is associated with a vibrational mode that the O\(_{\text{nuC}}\) attacks the phosphor with the leaving group (methoxide, CH\(_3\)O\(^-\)) departing simultaneously. This indicates that this is a concerted step. In our optimization, the transition state was observed to be converged in a quite flat region. The P-O\(_{\text{nuC}}\) and P-O\(_{\text{lg}}\) distances can easily change ca 0.2 Å without significant change of the energy. This finding is in line with the conclusion from the QM/MM calculations on the GTPase reaction of the RasGap system\(^{130}\), and is very important to the understanding of the nature of the enzyme-catalyzed phosphate ester hydrolysis. It turns out that, simultaneously with the P-O\(_{\text{nuC}}\) bond formation and the P-O\(_{\text{lg}}\) bond breaking, the proton of Ser102 is transferred to the O\(_{\text{MgH}}\), leading to one Mg-bound water molecule (H\(_2\)O\(_{\text{Mg}}\)). We could not locate an intermediate, where the proton of Ser102 has been delivered to the O\(_{\text{MgH}}\) but nucleophilic attack did not take place yet. This phosphoryl transfer step results in a phosphoenzyme intermediate (\(P\_{\text{S,me}}\)). The \(P\_{\text{S,me}}\) intermediate has been confirmed experimentally by the X-ray crystallography\(^{124}\). The energetic barrier of phosphoryl transfer is calculated to be 12.4 kcal/mol with the CPCM solvation (the potential energy surface is depicted in Figure 5.15). The \(P\_{\text{S,me}}\) intermediate is calculated to lie 4.8 kcal/mol higher than the \(R_{\text{me}}\) reactant complex.

To characterize the nature of phosphoryl transfer in the methyl phosphate hydrolysis, the potential energy surface is represented by a More O’Ferral Jencks (MFJ) plot (see the introduction of Section 4.3)\(^{58}\). We place the obtained stationary points into the MFJ plot in terms of their P-O\(_{\text{lg}}\) and P-O\(_{\text{nuC}}\) distances (Figure 5.13), and the nature of the transition state (\(T\_{\text{S,me}}\)) turns out to be associative. Therefore, it can be concluded that, in the AP-catalyzed methyl phosphate hydrolysis, the phosphoryl transfer proceeds along a concerted associative pathway. It should be mentioned here that many attempts have been made to find the dissociative pathways (concerted or step-wise) and the step-wise associative path, but without success.

It is of significance to check some geometrical changes that take place during the phosphoryl transfer. At \(T\_{\text{S,me}}\), most of distances between the Zn ions and the oxygen atoms of phosphate moiety, including Zn1-O\(_{\text{z1}}\), Zn2-O\(_{\text{lg}}\), and Zn2-O\(_{\text{z2}}\), become shorter relative to the \(R_{\text{me}}\) reactant (see Figure 5.13). This indicates that the two Zn ions play a role in stabilizing the transition state structure, thereby lowering the barrier.
for the phosphoryl transfer. This is consistent with the discussion in Section 4.4.3 and provide a strong support to the conclusion from the investigations of LFERs\textsuperscript{132} and crystallography\textsuperscript{125}, that the electrostatic interactions of phosphate with Zn ions are critical stabilizing factors in the transition state structure.

Another important parameter is the distance of Zn2 to the oxygen of the leaving group (Zn2-O\textsubscript{lg}). From a binding of 2.11 Å in the enzyme-substrate complex (R\textsubscript{me}), the distance is decreased to 1.96 Å in TS\textsubscript{me}, and finally 1.93 Å at the PSI\textsubscript{me} structure. As pointed out in Section 4.4.3, this demonstrates that another role of Zn2 is to stabilize the developing charge of the leaving group, also lowering the barrier of the phosphoryl transfer.

We also observe that the two hydrogen bond distances between the Arg166 and phosphate moiety change only little during the phosphoryl transfer (see Figure 5.13), suggesting that the Arg166 only makes a slight contribution to the catalysis. This may explain why mutation of this arginine residue does not inactivate the enzyme\textsuperscript{109,120}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure514.png}
\caption{Optimized structures of the stationary points along the phosphoryl transfer pathway in the p-nitrophenyl phosphate hydrolysis and the corresponding MFJ plot.}
\end{figure}
5.4.4 p-Nitrophenyl phosphate substrate

The phosphoryl transfer of the p-nitrophenyl phosphate hydrolysis has been investigated also. From the enzyme-substrate complex ($R_{pnp}$, see Figure 5.14), we have located the transition state ($T_{Spnp}$) for the nucleophile (Ser102 alkoxide) displacing the leaving group (p-nitrophenolate). The imaginary vibrational mode of the $T_{Spnp}$ indicates that, simultaneously with the P-O$_{nuc}$ bond formation, the P-O$_{lg}$ bond is breaking and the proton of Ser102 is transferred to the O$_{Me}H^-$. This step leads to a phosphoenzyme intermediate ($PS_{pnp}$) with the tetra-coordinated configuration at the phosphorus center. The barrier of this step is calculated to be 8.1 kcal/mol including the solvation, which is significantly lower than that in the methyl phosphate system (see Figure 5.15). In addition, it is calculated to be 23.3 kcal/mol downhill from the $T_{Spnp}$ to the $PS_{pnp}$ intermediate, making this step exothermic by 15.2 kcal/mol. This can be easily rationalized by considering the pK$_a$ values of the leaving groups. The p-nitrophynolate resulted in this step is a much better leaving group relative to the methoxide, due to its low pK$_a$ of 7.1$^{[133]}$. This can largely lower the energy of the phosphoseryl intermediate state and lead to a decrease of the reaction barrier. Furthermore, this coincides with the structural fact that the $T_{Spnp}$ is a significantly earlier transition state than the $T_{Sme}$, that is, the P-O$_{nuc}$ distance of $T_{Spnp}$ (2.27 Å) is much longer than that of $T_{Sme}$ (1.86 Å) while the P-O$_{lg}$ distance (2.07 Å) is much shorter (2.21 Å at $T_{Sme}$).

Also here, we place the obtained stationary points in the corresponding MFJ plot according to their P-O$_{lg}$ and P-O$_{nuc}$ distances (see Figure 5.14). It can be observed that the phosphoryl transfer in the p-nitrophenyl phosphate system also follows a concerted associative pathway, like the methyl phosphate system. Particular attention can be paid to the sum of P-O$_{lg}$ and P-O$_{nuc}$ distances at the transition state (P-O$_{lg}$ + P-O$_{nuc}$), since it can reflect the important chemical character of the phosphoryl transfer. It can be seen that the “P-O$_{lg}$ + P-O$_{nuc}$” value of p-nitrophenyl phosphate system (4.34 Å, see Figure 5.14) is larger than that of methyl phosphate system (4.07 Å, Figure 5.13). This indicates that the nature of the phosphoryl transfer tends to transform towards dissociative direction with the pK$_a$ of the leaving group decreasing, although it is associative for both methyl and p-nitrophenyl phosphate systems. This conclusion is consistent with the investigation by Warshel et al$^{[130]}$, who performed the systematic calculations on the potential surfaces of the reactions of a series of phosphate monoesters with different leaving groups.
5. Applications

5.4.5 Hydrolysis of phosphoseryl intermediate

Once the phosphoryl transfer is accomplished, the resulting methoxide (or p-nitrophenolate) may be protonated to release alcohol (or p-nitrophenol) product via two possible modes. One is to react with a water molecule activated by Zn2 ion and generate a Zn2-stabilized hydroxide (OZn2H−)\(^{[125]}\). Another possibility is that the methoxide is directly expelled into the solution and gets protonated there. For p-nitrophenolate, the protonation is possibly not necessary at all. Also, the OZn2H−, which acts as the nucleophile in the hydrolysis of the phosphoseryl intermediate, may come from the solution directly. No matter which mode is employed, the formation of Zn2-stabilized hydroxide is supposed to be a rapid process\(^{[108]}\). After the alcohol (or p-nitrophenol) product becomes free, the OZn2H− performs the second nucleophilic

---

**Figure 5.15** Calculated potential energy profiles for the AP-catalyzed reactions. (Methyl) Methyl phosphate system. (p-Nitrophenyl) p-Nitrophenyl phosphate system. (Cluster) B3LYP/6-311+G(2d,2p) energies with zero-point energies included. (CPCM) CPCM solvation effects added with \(\varepsilon = 4\).
5. Applications

...attack on the phosphorus center to replace the Ser102 alkoxide. Of course, a concerted process, where a water molecule performs the nucleophilic attack concomitant with water proton transfer to the methoxide (or p-nitrophenolate), can not be excluded.

To estimate the energetics of the protonation of methoxide (and p-nitrophenolate) bound by the AP active site, one water molecule was added into the phosphoseryl intermediates (PSIme and PSIpnp) to optimize the phosphoenzyme-alkoxide-water complexes (labeled by PSIme-W and PSIpnp-W, respectively) and the corresponding phosphoenzyme-alcohol-hydroxide complexes (PSIOH-ME and PSIOH-PNP). They correspond to the species before and after the protonation of the methoxide (and p-nitrophenolate), respectively. In addition, the phosphoenzyme-hydroxide complex without alcohol (or p-nitrophenol) product bound is referred to as PSIOH. The protonation of the methyl phosphate system, i.e. PSIme + H2O (in solution) → PSIOH + Methanol (in solution), makes the energy go down significantly (-8.0 kcal/mol with the solvation included, see Figure 5.15). In the optimization of phosphoenzyme-nitrophenol-hydroxide complex (PSIOH-PNP), the proton of nitrophenol is, however, always shuttled back to the OZn2H-, leading to the phosphoenzyme-nitrophenolate-water complex (PSIpnp-W). This possibly indicates that no activation (protonation) of the leaving group is necessary for the p-nitrophenyl phosphate system. It is thus postulated that the OZn2H- is introduced directly from the solution. By calculating individual species, the displacement of PSIpnp + OH- (in solution) → PSIOH + p-nitrophenolate (in solution) was predicted to be largely exothermic by 11.1 kcal/mol with the CPCM solvation (see Figure 5.15). However, the uncertainty in this value is quite large, because it involves solvation of a hydroxide ion using implicit solvent model.

After the alcohol product (or p-nitrophenolate) is released, the hydrolysis of phosphoseryl intermediate (PSI) takes place, i.e. the Zn2-bound hydroxide (OZn2H-) performs the nucleophilic attack on the phosphorus center to replace the Ser102 alkoxide. For all phosphate monoester substrates, this displacement step is entirely identical as the nucleophile (OZn2H-) and the leaving group (Ser102 alkoxide) are the same. The structure of phosphoenzyme-hydroxide complex (PSIOH) has been optimized and is displayed in Figure 5.16. It can be observed that, at PSIOH, the Zn2 ion plays a role in stabilizing and orienting the hydroxide in the optimal position to attack the phosphorus center. From PSIOH, the transition state for the hydrolysis of PSI (referred to as TSbhyd) has been optimized and is shown in Figure 5.16. It turns out that, when the P-O_nuc bond is formed, the P-O_{lg} bond is breaking and the proton of H2O_Mg is delivered back to the Ser102, leading to an inorganic phosphate (Pi) and the regenerated Ser102 residue and Mg-bound hydroxide (OMgH-). The optimized geometry of the resulting enzyme-product complex (Prod) is also given in Figure 5.16. The barrier and reaction energy of this step are calculated to be 16.0 and 5.9 kcal/mol including the solvation, respectively, (see Figure 5.15).
Figure 5.16 Optimized structures of the stationary points in the hydrolysis of phosphoseryl intermediate and the corresponding MFJ plot.

We found that the transition state of the hydrolysis of PSI was also converged in a flat region. The shift of ca 0.2 Å in the P-O$_{\text{nuc}}$ and P-O$_{\text{lg}}$ distances does not change the energy of the transition state significantly. This result is quite similar to the phosphoryl transfer in the methyl phosphate system, and the similarity may be attributed to the close pK$_a$ values between hydroxide and methoxide. The location of the stationary points in the corresponding MFJ plot (see Figure 5.16) shows that the hydrolysis of PSI also proceeds through a concerted associative path.

5.4.6 Conclusions

The calculations presented here confirm that the AP reaction employs a “ping-pong” mechanism, involving two displacement steps, i.e. the phosphoryl transfer and the hydrolysis of phosphoseryl intermediate (PSI). In this mechanism, a covalent phosphoseryl intermediate is formed and the configuration at phosphorus center is retained in the catalytic cycle. The two displacements both proceed through the concerted associative pathway for both methyl and p-nitrophenyl phosphates.
Furthermore, it has been observed that the transition states in the methyl phosphate hydrolysis are converged in a quite flat region, a feature that could be of significance for the interpretation of the LFER results.

The two zinc ions have been revealed by our calculations to stabilize the transition state structures, consequently lowering the barriers of phosphoryl transfer and hydrolysis of PSI. Zn2 even plays another role in stabilizing the developing charge at the leaving group, facilitating the phosphoryl transfer step.
Density functional theory studies of reaction mechanisms of a number of di-zinc enzymes have been presented in this thesis. The target enzymes are: phosphotriesterase (PTE), aminopeptidase from *Aeromonas proteolytica* (AAP), glyoxalase II (GlxII), and alkaline phosphatase (AP). Models of the active sites of these enzymes were constructed and the hydrolyses of different types of substrates were examined, including phosphate esters and the substrates containing carbonyl group.

Our calculations have demonstrated that the bridging hydroxide is capable of acting as the nucleophile in the hydrolysis. PTE, AAP, and GlxII all employ the bridging hydroxide as the direct nucleophile. The zinc ions provide the main catalytic power by stabilizing the charge developing during the reaction and consequently lowering the barriers. In the cases of GlxII and AP, one of zinc ions also contributes to the catalysis by stabilizing the leaving group. These features perfectly satisfy the two requisites for the hydrolysis, i.e. (i) sufficient nucleophilicity and (ii) stabilization of charge (mentioned in Chapter 4).

It has furthermore been demonstrated that the nucleophile mechanism is preferable for Zn/Zn-PTE reaction, i.e. the bridging hydroxide acts not as a base but as a direct nucleophile to perform the attack on the substrate.

The associative mechanism has been detected for both the PTE and AP reactions. The former uses a step-wise associative pathway via a penta-coordinated intermediate, while the latter proceeds through a concerted associative path.

Finally, with PTE as the test case, a systematic analysis has been performed to assess the adequacy of some of the technical approximations frequently used in the modeling of enzyme reactions. These include the choice of basis set for geometry optimizations and energy evaluation, the choice of dielectric constant to model the enzyme surrounding, and the effects of locking the centers of truncation. This evaluation and other results of this thesis offer an effective verification of the usefulness and powerfulness of quantum chemical active-site modeling in the exploration of enzyme reaction mechanisms and in the characterization of the transition states involved.
References

[10] Hopmann, K. H.; Himo, F.
[14] Siegbahn, P.E.M.
[15] Himo, F.; Siegbahn, P.E.M.
[16] Benkovic, S.; Hammes-Schiffer, S.
[18] Eisenmesser, E.Z.; Bosco, D.A.; Akke, M.; Kern, D.
[19] Benkovic, S.J.; Cannon, W.R.
[21] Bugg, T.
[22] Smith, S.
[24] Xiang, Y.; Liu, H.; Olsson, M. H. M.
### References


[51] (a) Kaminskaia, N. V.; He, C.; Lippard, S. J.

(a) Klahn, M.; Rosta, E.; Warshel, A.

(b) Williams, A.

(c) Jencks, W. P.

(d) Cleland, W. W.; Hengge, A. C.

Klabunde, T.; Sträter, N.; Fröhlich, R.; Witzel, H.; Krebs, B.


Klabunde, T.; Sträter, N.; Fröhlich, R.; Witzel, H.; Krebs, B.


(a) Weston, J. *Chem. Rev.* 2005, 105, 2151. (b) Seibert, C.M.; Rauschel, F.M. *Biochemistry* 2005,
References

44, 6383.


